PROBABILITY COMPREHENSIVE EXAM SPRING 2015

Problem 1. *Assume X is a symmetric random variable such that $\mathbb{E}[X^2] = 1$ and $\mathbb{E}[X^4] = 2$. Show that*

$$\mathbb{P}(X \geq 1) \leq \frac{14}{27}.$$

Solution. If we try a brute force Chebyshev inequality we obtain a trivial bound, namely that $\mathbb{P}(X \geq 1) \leq 1$, something not very useful. The idea is to use the Chebyshev's inequality in a cleverer way. Take first a positive $t \geq 0$ and then write

$$\mathbb{P}(X \geq 1) = \mathbb{P}(X + t \geq 1 + t) \leq \mathbb{P}((X + t)^4 \geq (1 + t)^4).$$

Then use Chebyshev's inequality to continue with

$$\mathbb{P}(X+t \geq 1+t) \leq \mathbb{P}((X+t)^4 \geq (1+t)^4) \leq \frac{\mathbb{E}[(X+t)^4]}{(1+t)^4} = \frac{\mathbb{E}[X^4] + 4t\mathbb{E}[X^3] + 6t^2\mathbb{E}[X^2] + 4t^3\mathbb{E}[X] + t^4}{(1+t)^4}.$$

Since the variable is symmetric, $\mathbb{E}[X] = 0$ and also $\mathbb{E}[X^3] = 0$. Thus,

$$\mathbb{P}(X \geq 1) \leq \frac{2 + 6t^2 + t^4}{(t+1)^4}.$$

Now we want to take the best possible choice for t and for that matter we need to find the minimum value of

$$f(t) = \frac{2 + 6t^2 + t^4}{(t+1)^4}$$

Taking the derivative, we get

$$f'(t) = \frac{4\,(t^3 - 3t^2 + 3t - 2)}{(t+1)^5} = \frac{4(t-2)(t^2 - t + 1)}{(t+1)^5}.$$

Therefore, $t = 2$ is a critical point and it is also a minimum point. This means

$$\mathbb{P}(X \geq 1) \leq f(2) = \frac{14}{27}.$$

□

Problem 2. *Assume that $X_1, X_2, \ldots, X_n, \ldots$ is a sequence of iid random variables such that for some $\alpha < 1/2$,*

$$\frac{X_1 + X_2 + \cdots + X_n}{n^\alpha} \xrightarrow[n\to\infty]{a.s.} m$$

for some real number m (and the convergence is in the almost sure sense). Show that almost surely $X_i = 0$.

Solution. We prove in the first place that X_i are integrable. That is a standard application of Borel-Cantelli. Denote $S_n = X_1 + X_2 + \cdots + X_n$. Then

$$\frac{X_n}{n^\alpha} = \frac{S_n}{n^\alpha} - \frac{S_{n-1}}{n^\alpha} \xrightarrow[n\to\infty]{a.s.} 0.$$

Now we show that the the variable is square integrable. To do this we start with the fact that

$$\frac{X_n^2}{n^{2\alpha}} \xrightarrow[n\to\infty]{a.s.} 0.$$

thus, we have by Borel-Cantelli and the fact that X_n are independent that

$$\sum_{n\geq 1}\mathbb{P}\left(\frac{X_n^2}{n^{2\alpha}}\geq\epsilon\right)<\infty.$$

Indeed, the divergence of the series and independence would mean that $\mathbb{P}\left(\frac{X_n^2}{n^{2\alpha}}\geq\epsilon\, i.o.\right)=1$ and this would contradict the convergence of the $\frac{X_n}{n^\alpha}$ to 0.

Now, we have that

$$\sum_{n\geq 1}\mathbb{P}\left(\frac{X_n^2}{n^{2\alpha}}\geq 1\right)=\sum_{n\geq 1}\mathbb{P}\left(\frac{X_1^2}{n^{2\alpha}}\geq 1\right)=\sum_{n\geq 1}\mathbb{P}\left(X_1^2\geq n^{2\alpha}\right)\leq\sum_{n\geq 1}\mathbb{P}\left(X_1^2\geq n\right)<\infty$$

This last part implies that the variable X_1^2 is integrable. In particular X_1 is also integrable.

On the other hand, since $\alpha<1/2$, we can also conclude that

$$\frac{X_1+X_2+\cdots+X_n}{n}\xrightarrow[n\to\infty]{a.s.}0$$

and thus, by the Strong Law of large numbers, the mean must be 0.

Now if we assume that $\mathbb{E}[X_1^2]=\sigma^2>0$, then the central limit theorem gives us that

$$\frac{X_1+X_2+\cdots+X_n}{n^{1/2}}\Longrightarrow N\left(0,\sigma^2\right).$$

On the other hand from the given condition ($\alpha<1/2$) we also get

$$\frac{X_1+X_2+\cdots+X_n}{n^{1/2}}\xrightarrow[n\to\infty]{a.s.}0$$

and since a.s. convergence implies the weak convergence we get that $N(0,\sigma^2)=0$, which is a contradiction.

Thus $\mathbb{E}[X_1^2]=0$ and this means that $X_1=0$ almost surely. $\square$

Problem 3. *Assume that (X,Y) is a joint normal vector with $\mathbb{E}[X]=\mathbb{E}[Y]=0$. Show that*

$$\mathbb{E}[X^2Y^2]\geq\mathbb{E}[X^2]\mathbb{E}[Y^2]$$

with equality if and only if X and Y are independent.

Solution. For simplicity we may assume that the variance of X and Y are both equal to 1. Then the covariance matrix can be written as

$$C=\begin{bmatrix}1&\rho\\ \rho&1\end{bmatrix}.$$

We learn from this that the characteristic function of (X,Y) is determined by

$$f_{X,Y}(\xi,\eta)=\mathbb{E}[e^{i\xi X+i\eta Y}]=\exp\left(-\frac{\xi^2+\eta^2+2\rho\xi\eta}{2}\right)$$

Then, the integral $\mathbb{E}[X^2Y^2]$ can be computed as the derivative of the characteristic function as

$$E[X^2Y^2]=\frac{1}{i^4}\frac{\partial^4}{\partial\xi^2\partial\eta^2}f_{X,Y}(\xi,\eta)\bigg|_{\xi=\eta=0}=\frac{\partial^4}{\partial\xi^2\partial\eta^2}f_{X,Y}(\xi,\eta)\bigg|_{\xi=\eta=0}.$$

This last part can be computed from the characteristic function in the form

$$\frac{\partial^4}{\partial\xi^2\partial\eta^2}f_{X,Y}(\xi,\eta)\bigg|_{\xi=\eta=0}=\frac{\partial^4}{\partial\xi^2\partial\eta^2}\exp\left(-\frac{\xi^2+\eta^2+2\rho\xi\eta}{2}\right)\bigg|_{\xi=\eta=0}=1+\frac{\rho^2}{2}.$$

Thus
$$\mathbb{E}[X^2Y^2] = 1 + \frac{\rho^2}{2} \geq 1 = \mathbb{E}[X^2]\mathbb{E}[Y^2]$$
with equality if and only if $\rho = 0$ which is the same as saying that X and Y are uncorrelated which because of the fact that (X, Y) is a normal vector, is equivalent to saying that X and Y are independent. □

Problem 4. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and $(\mathcal{F})_{n\geq 0}$ is a filtration on it. Show that if $(M_n)_{n\geq 0}$ is a martingale such that $(M_n^4)_{n\geq 0}$ is also a martingale, then almost surely $M_n = M_0$ for any $n \geq 0$.*

Solution. Since M_n^4 is a martingale, we write using the properties of the conditional expectation and the fact that M_n is $\mathcal{F}_n$-measurable,
$$\begin{aligned}
\mathbb{E}[M_{n+1}^4|\mathcal{F}_n] &= \mathbb{E}[(M_{n+1} - M_n + M_n)^4|\mathcal{F}_n] \\
&= \mathbb{E}[(M_{n+1} - M_n)^4|\mathcal{F}_n] + \mathbb{E}[4(M_{n+1} - M_n)^3 M_n|\mathcal{F}_n] + 6\mathbb{E}[(M_{n+1} - M_n)^2 M_n^2|\mathcal{F}_n] \\
&\quad + 4\mathbb{E}[(M_{n+1} - M_n)M_n^3|\mathcal{F}_n] + \mathbb{E}[M_n^4|\mathcal{F}_n] \\
&= \mathbb{E}[(M_{n+1} - M_n)^2((M_{n+1} - M_n)^2 + 4(M_{n+1} - M_n)M_n + 6M_n^2)|\mathcal{F}_n] \\
&\quad + 4M_n^3\mathbb{E}[M_{n+1} - M_n|\mathcal{F}_n] + M_n^4.
\end{aligned}$$
Now, from the fact that M_n is a martingale, $\mathbb{E}[M_{n+1} - M_n|\mathcal{F}_n] = 0$ and thus we obtain that
$$\mathbb{E}[(M_{n+1} - M_n)^2((M_{n+1} - M_n)^2 + 4(M_{n+1} - M_n)M_n + 6M_n^2)|\mathcal{F}_n] = 0.$$
Simplifying this a little bit we arrive at
$$\mathbb{E}[(M_{n+1} - M_n)^2((M_{n+1} - M_n + 2M_n)^2 + 2M_n^2)|\mathcal{F}_n] = 0$$
Taking now the expectation we get
$$\mathbb{E}[(M_{n+1} - M_n)^2((M_{n+1} + M_n)^2 + 2M_n^2)] = 0$$
which yields that either $M_{n+1} = M_n$ almost surely, or $(M_{n+1} + M_n)^2 + 2M_n^2 = 0$ which again implies that $M_{n+1} = M_n$ almost surely. □

Problem 5. *For a sequence $X_1, X_2, \ldots, X_n, \ldots$ we know that*
$$\sum_{n=1}^{\infty} n\mathbb{E}[|X_n|] < \infty.$$
Show that the sequence $Y_n = X_n + X_{n+1} + \cdots + X_{10n}$, converges almost surely and in L^1 to 0.

Solution. The convergence in L^1, follows from the fact that
$$\mathbb{E}[|Y_n|] \leq \sum_{k=n+1}^{10n} \mathbb{E}[|X_k|] \leq \sum_{k=n+1}^{\infty} k\mathbb{E}[|X_k|] \xrightarrow[n\to\infty]{} 0$$

For the almost sure convergence, notice that summing over n the inequalities $\mathbb{E}[|Y_n|] \leq \sum_{k=n+1}^{\infty} \mathbb{E}[|X_k|]$, we then get
$$\sum_{n=1}^{\infty} \mathbb{E}[|Y_n|] \leq \sum_{n=1}^{\infty} (n+1)\mathbb{E}[|X_n|] \leq 2\sum_{n=1}^{\infty} n\mathbb{E}[|X_n|] < \infty$$
Therefore the almost sure convergence follows from Chebyshev's inequality and Borel-Cantelli lemma and the observation that
$$\sum_{n=1}^{\infty} \mathbb{P}(|Y_n| \geq \epsilon) \leq \frac{1}{\epsilon}\sum_{n=1}^{\infty} \mathbb{E}[|Y_n|] < \infty.$$
□

Problem 6. *Assume $\{X_n\}_{n\geq 1}$ is a sequence of iid random variables with mean 0 and variance 1. Show that*

$$Y_n = \frac{\sqrt{n}X_1 + \sqrt{n-1}X_2 + \sqrt{n-2}X_3 + \cdots + X_n}{n}$$

converges weakly (in distribution) to a normal $N(0, 1/2)$.

Solution. It is enough to show that the characteristic functions converge. Now, the characteristic function of Y_n is computed as

$$f_{Y_n}(\xi) = f_{X_1}(\sqrt{n}\xi/n) f_{X_2}(\sqrt{n-1}\xi/n) \dots f_{X_n}(\xi/n).$$

Since $X_1, X_2, \dots, X_n$ are identically distributed, we denote by $f(\xi) = f_{X_i}(\xi)$ for any i and then continue with

$$f_{Y_n}(\xi) = f(\sqrt{n}\xi/n) f(\sqrt{n-1}\xi/n) \dots f(\xi/n).$$

Since X_1 has mean 0 and variance 1, $f(\xi) = 1 - \xi^2/2 + o(\xi^2)$ which written in exponential form for small enough ξ, gives $f(\xi) = e^{-\frac{\xi^2}{2} + o(\xi^2)}$. From this we deduce that for large n and fixed ξ

$$f(\sqrt{n}\xi/n) f(\sqrt{n-1}\xi/n) \dots f(\xi/n) = \exp(-\frac{n\xi^2}{2n^2} - \frac{(n-1)\xi^2}{2n^2} - \dots - \frac{\xi^2}{2n^2} - o(1)) = \exp(-\frac{\xi^2}{4} + o(1))$$

from which we get that

$$f_{Y_n}(\xi) \xrightarrow{n\to\infty} \exp(-\frac{\xi^2}{4}) = f_{N(0,1/2)}(\xi)$$

and this completes the proof. $\square$

Problem 7. *Assume that $\{U_n\}_{n\geq 1}$ is a sequence of iid uniform random variables on $[0,1]$. Let $V_n = \max\{U_1, U_2^2, \dots, U_n^n\}$. Show that $(1 - V_n)\ln(n)$ converges weakly (in distribution) to an exponential random variable with parameter 1.*

Solution. To do this we will compute the cumulative function of $W_n = (1 - V_n)\ln(n)$. Take a positive x and let n be large enough. Now, using independence,

$$\begin{aligned}
1 - F_{W_n}(x) &= \mathbb{P}(W_n > x) = \mathbb{P}\left(V_n < 1 - \frac{x}{\ln(n)}\right) \\
&= \mathbb{P}\left(U_1 < 1 - \frac{x}{\ln(n)}, U_2^2 < 1 - \frac{x}{\ln(n)}, \dots, U_n^n < 1 - \frac{x}{\ln(n)}\right) \\
&= \mathbb{P}\left(U_1 < 1 - \frac{x}{\ln(n)}\right) \mathbb{P}\left(U_2^2 < 1 - \frac{x}{\ln(n)}\right), \dots, \mathbb{P}\left(U_n^n < 1 - \frac{x}{\ln(n)}\right) \\
&= \mathbb{P}\left(U_1 < 1 - \frac{x}{\ln(n)}\right) \mathbb{P}\left(U_2 < \left(1 - \frac{x}{\ln(n)}\right)^{1/2}\right), \dots, \mathbb{P}\left(U_n < \left(1 - \frac{x}{\ln(n)}\right)^{1/n}\right) \\
&= \left(1 - \frac{x}{\ln(n)}\right)\left(1 - \frac{x}{\ln(n)}\right)^{1/2} \dots \left(1 - \frac{x}{\ln(n)}\right)^{1/n} \\
&= \left(1 - \frac{x}{\ln(n)}\right)^{1+1/2+\dots+1/n} = \left(\left(1 - \frac{x}{\ln(n)}\right)^{\ln(n)}\right)^{\frac{1+1/2+\dots+1/n}{\ln(n)}}.
\end{aligned}$$

Finally, since $(1 - x/\ln(n))^{\ln(n)}$ converges to e^{-x} as n tends to infinity and

$$1 + 1/2 + \dots + \frac{1}{n} - (1 + 1/n) < \ln(n) < 1 + 1/2 + \dots + 1/n$$

we obtain that
$$\frac{1+1/2+\cdots+1/n}{\ln(n)} \xrightarrow{n\to\infty} 1$$
and thus
$$1-F_{W_n}(x) \xrightarrow{n\to\infty} e^{-x} = 1-F_Z(x)$$
where Z is an exponential random variable with parameter 1.

$\square$

Problem 8. *Let $\{X_n\}_{n\geq 1}$ be an iid sequence of positive random variables such that $E[X_1]<\infty$. Let*
$$N_t = \sup\{n : X_1+X_2+\cdots+X_n \leq t\}.$$
Show that
$$\frac{N_t}{t} \xrightarrow[t\to\infty]{a.s.} \frac{1}{\mathbb{E}[X_1]}$$
where the convergence is in almost sure sense.

Solution. Let $S_n = X_1+X_2+\cdots+X_n$. From the Law of Large numbers, we learn that S_n converges to $+\infty$ as n tends to ∞. In particular N_t is almost surely finite and $N(t)$ tends to infinity with t. Now for any t we obviously have
$$S_{N(t)} \leq t \leq S_{N(t)+1}$$
thus,
$$\frac{N(t)}{S_{N(t)+1}} \leq \frac{N(t)}{t} \leq \frac{N(t)}{S_{N(t)}}.$$
Now, from the Law of large numbers we have that
$$\frac{S_n}{n} \xrightarrow[n\to\infty]{a.s.} \mathbb{E}[X_1]$$
which in turn combined with the fact that N_t converges to infinity with t concludes the proof.

$\square$

Probability Comprehensive Exam
Aug 24, 2016

Student Number:

Instructions: Complete 5 of the 8 problems, and **circle** their numbers below – the uncircled problems will **not** be graded.

1 2 3 4 5 6 7 8

Write **only on the front side** of the solution pages. A **complete solution** of a problem is preferable to partial progress on several problems.

1. Suppose (X_n) is a sequence of random vectors such that for some sigma-algebra $\mathcal{F}$, one has X_n and $\mathcal{F}$ independent for all n. If $X_n \to X$ almost surely, show that X and $\mathcal{F}$ are independent.

Solution: Let $A \in \mathcal{F}$. If $\mathbb{P}(A) \in \{0, 1\}$, then A is independent of all events, in particular of any event in the sigma-algebra generated by X. So we may assume that $\mathbb{P}(A) \in (0, 1)$ and prove that A is independent of X.

For a random vector Z, let ϕ_Z be its characteristic function. Define the conditional measure $\mathbb{P}_A(\cdot) = \mathbb{P}(\cdot \mid A)$ and denote the corresponding characteristic function of a random variable Z by ϕ_Z^A. Then since $X_n \to X$ almost surely, the convergence also occurs in distribution relative to $\mathbb{P}$, and so $\phi_{X_n} \to \phi_X$ pointwise. By independence between X_n and A, one has

$$\phi_{X_n}^A(t) = \frac{\mathbb{E}e^{it\cdot X_n}\mathbf{1}_A}{\mathbb{P}(A)} = \phi_{X_n}(t),$$

so $\phi_{X_n}^A \to \phi_X$ pointwise as well. Since $X_n \to X$ almost surely, also for any t,

$$e^{it\cdot X_n}\mathbf{1}_A \to e^{it\cdot X}\mathbf{1}_A \text{ almost surely.}$$

By the bounded convergence theorem,

$$\phi_{X_n}^A(t) = \frac{\mathbb{E}e^{it\cdot X_n}\mathbf{1}_A}{\mathbb{P}(A)} \to \frac{\mathbb{E}e^{it\cdot X}\mathbf{1}_A}{\mathbb{P}(A)} = \phi_X^A(t).$$

This implies $\phi_X(t) = \phi_X^A(t)$, so since characteristic functions determine a distribution, the distribution of X under $\mathbb{P}$ is the same as that under $\mathbb{P}_A$. In other words, for $B \subset \mathbb{R}^n$ Borel,

$$\mathbb{P}(X \in B) = \mathbb{P}_A(X \in B) = \frac{\mathbb{P}(X \in B,\ A)}{\mathbb{P}(A)},$$

or X and A are independent. This implies X and $\mathcal{F}$ are independent.

2. Let X be a random variable with continuous density function f and $f(0) > 0$. Let Y be a random variable with

$$Y = \begin{cases} \frac{1}{X} & \text{if } X > 0 \\ 0 & \text{otherwise} \end{cases}.$$

and $Y_1, Y_2, \ldots$ be i.i.d. with distribution equal to that of Y. What is the value of the almost sure limit

$$\lim_n \frac{Y_1 + \cdots + Y_n}{n}?$$

Solution: First compute the probability for $y > 0$

$$\mathbb{P}(Y \geq y) = \mathbb{P}(X \in (0, 1/y]).$$

Choose $x_0 > 0$ small enough so that $f(x) \geq f(0)/2$ for $x \in [0, x_0]$. Then for $y > 1/x_0$, one has

$$\mathbb{P}(Y \geq y) = \int_0^{1/y} f(x)\ \mathrm{d}x \geq \frac{f(0)}{2y}.$$

Therefore

$$\begin{aligned}\mathbb{E}Y = \int_0^\infty \mathbb{P}(Y \geq y)\ \mathrm{d}y &\geq \int_{1/x_0}^\infty \mathbb{P}(Y \geq y)\ \mathrm{d}y \\ &\geq \int_{1/x_0}^\infty \frac{f(0)}{2y}\ \mathrm{d}y = \infty.\end{aligned}$$

Now since the Y_i's have infinite mean and are positive, we can show that $(Y_1 + \cdots + Y_n)/n \to \infty$ almost surely. To do so, pick any $M > 0$ and define

$$Y_i^{(M)} = \min\{Y_i, M\}.$$

By the strong law of large numbers,

$$\frac{Y_1^{(M)} + \cdots + Y_n^{(M)}}{n} \to \mathbb{E}Y_1^{(M)} \text{ almost surely.}$$

Therefore, for any $M > 0$, almost surely

$$\liminf_{n\to\infty} \frac{Y_1 + \cdots + Y_n}{n} \geq \lim_n \frac{Y_1^{(M)} + \cdots + Y_n^{(M)}}{n} = \mathbb{E}Y_1^{(M)}.$$

The event on which this inequality holds we denote by A_M. Then on $\cap_{M\in\mathbb{N}} A_M$ (which has probability one), we have

$$\liminf_{n\to\infty} \frac{Y_1 + \cdots + Y_n}{n} \geq \sup_{M\in\mathbb{N}} \mathbb{E}Y_1^{(M)} = \lim_{M\to\infty} \mathbb{E}Y_1^{(M)} = \infty,$$

where the last equation holds by the monotone convergence theorem.

3. Let $X_1, X_2, \ldots$ be i.i.d. with

$$\mathbb{P}(X_1 = 1) = 1/2 = \mathbb{P}(X_1 = -1).$$

Let C be the set of factorials:

$$C = \{k! : k \in \mathbb{N}\}.$$

Show that

$$\lim_n \mathbb{P}(X_1 + \cdots + X_n \in C) = 0.$$

Hint. You may want to start by covering part of $[0,\infty)$ by small intervals. Let $\epsilon > 0$ and $I > 0$, consider intervals $I_1, \ldots, I_I$, where $I_i = [(i-1)\epsilon, i\epsilon)$, and show that for large n, at most two of the sets $I_i \cap (C/\sqrt{n})$ are nonempty.

Solution: Fix $\delta > 0, \epsilon > 0$ and consider the sets

$$I_1 = [0,\epsilon),\ \ I_2 = [\epsilon, 2\epsilon),$$

and generally $I_i = [(i-1)\epsilon, i\epsilon)$. Write $C/\sqrt{n} = \{k!/\sqrt{n} : k \in \mathbb{N}\}$ and for $I > 0$ fixed, estimate

$$\limsup_n \mathbb{P}(S_n \in C) \le \limsup_n \Bigg[\sum_{i=1}^{I} \mathbb{P}(S_n/\sqrt{n} \in (C/\sqrt{n}) \cap I_i) + \mathbb{P}(S_n/\sqrt{n} \ge I\epsilon) \Bigg].$$

We claim that for large n, at most one of the sets $(C/\sqrt{n}) \cap I_i$ is nonempty, for $i = 2, \ldots, I$. Indeed, assuming at least one is nonempty, let $i_0 \ge 2$ be the minimal index i such that $(C/\sqrt{n}) \cap I_i$ is nonempty. Then there is $\ell \in \mathbb{N}$ such that

$$\ell!/\sqrt{n} \in [(i_0 - 1)\epsilon, i_0\epsilon).$$

Note that since $\ell! \ge \epsilon\sqrt{n}$, we must have for n large

$$\ell \ge I.$$

For such large n,

$$(\ell+1)!/\sqrt{n} = (\ell+1) \cdot \frac{\ell!}{\sqrt{n}} \ge \ell\epsilon \ge I\epsilon,$$

so all of the sets $(C/\sqrt{n}) \cap I_i$ for $i = i_0 + 1, \ldots, I$ are empty.

Putting this together, we get

$$\limsup_n \mathbb{P}(S_n \in C) \le 3 \limsup_n \max \left\{ \mathbb{P}(S_n/\sqrt{n} \ge I\epsilon), \max_{i \in [1,I]} \mathbb{P}(S_n/\sqrt{n} \in I_i) \right\}.$$

By the CLT, if X is a standard normal random variable, this converges to

$$3 \limsup_n \max \left\{ \mathbb{P}(X \ge I\epsilon), \max_{i \in [1,I]} \mathbb{P}(|X| \in I_i) \right\}.$$

Since the distribution of X has no atoms, we may choose ϵ so small and I so large that this is at most δ.

4. Let (X, Y) be a normal vector in $\mathbb{R}^2$ with mean zero and covariance matrix Σ, where

$$\Sigma = \begin{pmatrix} 5 & 1 \\ 1 & 10 \end{pmatrix}.$$

Find $\mathbb{E}X^2Y^2$.

Solution: If (X_1, X_2) is a standard Gaussian vector, then if A is an arbitrary 2×2 real matrix,

$$A \cdot \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} =: \begin{pmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix} \cdot \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$$

is an arbitrary mean zero Gaussian vector (Y_1, Y_2). The covariance matrix Σ is given by

$$\Sigma_{i,j} = \mathbb{E}Y_iY_j = \sum_{k,\ell=1}^{2} \mathbb{E}A_{k,i}X_kA_{\ell,j}X_\ell = \sum_{k=1}^{2} A_{k,i}A_{k,j} = (A^TA)_{i,j}.$$

Solving in our case for A in terms of the covariance matrix Σ, we find that

$$X = X_1 + 2X_2,\ Y = 3X_1 - X_2,$$

where (X, Y) is in the statement of the problem, and (X_1, X_2) is a standard normal vector.

Now we can compute the expected value as

$$\begin{aligned} &\mathbb{E}(X_1 + 2X_2)^2(3X_1 - X_2)^2 \\ &= \mathbb{E}((X_1^2 + 4X_1X_2 + 4X_2^2)(9X_1^2 - 6X_1X_2 + X_2^2)) \\ &= \mathbb{E}(9X_1^4 + 30X_1^3X_2 + 13X_1^2X_2^2 - 20X_1X_2^3 + 4X_2^4). \end{aligned}$$

Using the i.i.d. assumption with symmetry, we obtain

$$13\mathbb{E}X_1^4 + 13(\mathbb{E}X_1^2)^2 = 13 \cdot 3 + 13 = 52.$$

5. Let ξ_1 and ξ_2 be independent random variables with characteristic functions $\varphi_1(u) = \frac{1-iu}{1+u^2}$ and $\varphi_2(u) = \frac{1+iu}{1+u^2}$ respectively. Find the probability that $\xi_1 + \xi_2$ takes values in $(3, +\infty)$.

Solution: Since ξ_1 and ξ_2 are independent, the characteristic function of $\xi_1 + \xi_2$ equals the product of their characteristic functions:

$$\varphi(u) = \varphi_1(u)\varphi_2(u) = \frac{1}{1 + u^2}.$$

Therefore, $\xi_1 + \xi_2$ has bilateral distribution with density

$$f(x) = \frac{1}{2} e^{-|x|},$$

and

$$\mathbb{P}(\xi_1 + \xi_2 > 3) = \frac{1}{2} \int_3^\infty e^{-x} dx = \frac{e^{-3}}{2}.$$

Answer: $\frac{e^{-3}}{2}$.

6. Let $\{A_n\}$ be an infinite collection of independent events. Suppose that $\mathbb{P}(A_n) < 1$ for every $n \geq 1$. Show that $\mathbb{P}(A_n\, i.o.) = 1$ if and only if $\mathbb{P}(\cup A_n) = 1$.

Solution: For any $N \geq 1$,

$$\begin{aligned} \mathbb{P}(A_1^c \cap \cdots \cap A_N^c)\mathbb{P}(\cup_{n>N} A_n) &= \mathbb{P}(A_1^c \cap \ldots \cap A_N^c \cap (\cup_{n>N} A_n)) \\ &= \mathbb{P}(A_1^c \cap \ldots \cap A_N^c \cap (\cup_{n\geq 1} A_n)). \end{aligned} \tag{1}$$

If $\mathbb{P}(\cup_{n\geq 1} A_n) = 1$, then (1) equals

$$\mathbb{P}(A_1^c \cap \cdots \cap A_N^c).$$

Since $\mathbb{P}(A_n) < 1$ for all n,

$$\mathbb{P}(A_1^c \cap \cdots \cap A_N^c) = \prod_{i=1}^N \mathbb{P}(A_i^c) > 0,$$

so we may divide both sides of

$$\mathbb{P}(A_1^c \cap \cdots \cap A_N^c)\mathbb{P}(\cup_{n>N} A_n) = \mathbb{P}(A_1^c \cap \cdots \cap A_N^c)$$

by $\mathbb{P}(A_1^c \cap \cdots \cap A_N^c)$, and we get

$$\mathbb{P}(\cup_{n>N} A_n) = 1.$$

This is true for all $N \geq 1$, hence

$$P(A_n i.o.) = P(\cap_N \cup_{n\geq N} A_n) = 1.$$

Conversely, if $P(A_n\, i.o.) = 1$, then since

$$\{A_n \text{ i.o.}\} = \cap_N \cup_{n\geq N} A_n \subset \cup_n A_n$$

we also have $\mathbb{P}(\cup_n A_n) = 1$.

7. Let X be a random variable taking values on the interval $[1, 2]$. Find sharp lower and upper estimates on the quantity $\mathbb{E}X \cdot \mathbb{E}\frac{1}{X}$. Provide an example of a random variable for which the lower estimate is attained. Provide an example of a random variable for which the upper estimate is attained.

   Hint. For the upper bound, justify and use the inequality
$$ab \leq \frac{1}{2}\left(\frac{a}{2} + b\right)^2.$$

Solution: To obtain the lower bound, we use Jensen's inequality:
$$\mathbb{E}X \cdot \mathbb{E}\frac{1}{X} \geq \mathbb{E}X \cdot \frac{1}{\mathbb{E}X} = 1.$$
The lower bound is attained when $X = c$ almost surely for some $c \in [1, 2]$.

Let $f(x)$ be the density of X. Let μ be the distribution of X. As for the upper bound, we apply the inequality $ab \leq \frac{1}{2}\left(\frac{a}{2} + b\right)^2$ with
$$a = \mathbb{E}X \text{ and } b = \mathbb{E}\frac{1}{X},$$
giving
$$\begin{aligned}
\mathbb{E}X \cdot \mathbb{E}\frac{1}{X} &= \int_1^2 x \, \mathrm{d}\mu(x) \cdot \int_1^2 \frac{1}{x} \, \mathrm{d}\mu(x) \\
&= 2\int_1^2 \frac{x}{2} \, \mathrm{d}\mu(x) \cdot \int_1^2 \frac{1}{x} \, \mathrm{d}\mu(x) \\
&\leq 2 \cdot \frac{1}{4}\left(\int_1^2 \frac{x}{2} f(x)dx + \int_1^2 \frac{1}{x} f(x)dx\right)^2 \\
&= \frac{1}{2}\left(\int_1^2 (\frac{x}{2} + \frac{1}{x}) f(x)dx\right)^2.
\end{aligned}$$
Here we used the inequality
$$ab \leq \frac{1}{4}(a+b)^2.$$
$$\mathbb{E}X \cdot \mathbb{E}\frac{1}{X} \leq \frac{1}{2}\left(\frac{1}{2}\mathbb{E}X + \mathbb{E}\frac{1}{X}\right)^2 = \frac{1}{2}\left(\mathbb{E}\left(\frac{X}{2} + \frac{1}{X}\right)\right)^2.$$
Observe that the maximum of $\frac{x}{2} + \frac{1}{x}$ over $[1, 2]$ is attained when $x = 2$ (or when $x = 1$), where the function takes value $\frac{3}{2}$. Therefore,
$$\mathbb{E}X \cdot \mathbb{E}\frac{1}{X} \leq \frac{1}{2}\left(\frac{3}{2}\right)^2 = \frac{9}{8}.$$

For an example attaining this value, let the random variable X_0 take values 1 and 2 with probability $\frac{1}{2}$ each. Then

$$\mathbb{E}X_0 \cdot \mathbb{E}\frac{1}{X_0} = \frac{3}{2} \cdot \frac{3}{4} = \frac{9}{8}.$$

8. Show that for a sequence of random variables X_n, one has $X_n \to X$ in probability if and only if
$$\mathbb{E}\left[e^{\min\{2,|X_n-X|\}} - 1\right] \to 0,$$
as $n \to \infty$.

Solution: First suppose that $X_n \to X$ in probability. Then also $|X_n - X| \to 0$ in probability and therefore in distribution. By Portmanteau's theorem, for any bounded continuous $f : \mathbb{R} \to \mathbb{R}$, one has $\mathbb{E}f(|X_n - X|) \to f(0)$. Applying this to the function $f(x) = e^{\min\{2,x\}} - 1$, we obtain the convergence in the problem.

and let $\epsilon > 0$. Chose n_0 large enough such that for every $n \geq n_0$,

$$\mathbb{P}\left(|X_n - X| \geq \min\left\{\frac{\epsilon}{2}, 1\right\}\right) < e^{-2}\frac{\epsilon}{2}.$$

But then, for all $n \geq n_0$,

$$\begin{aligned}
&\mathbb{E}e^{\min\{2,|X-X_n|\}} \\
&= \int_{|X_n-X|\geq\min\{\frac{\epsilon}{2},1\}} e^{\min\{2,|X-X_n|\}} \, \mathrm{d}\mathbb{P} + \int_{|X_n-X|<\min\{\frac{\epsilon}{2},1\}} e^{\min\{2,|X-X_n|\}} \, \mathrm{d}\mathbb{P} \\
&\leq e^2\mathbb{P}\left(|X_n - X| \geq \min\left\{\frac{\epsilon}{2}, 1\right\}\right) + \min\left\{\frac{\epsilon}{2}, 1\right\} \\
&< \epsilon.
\end{aligned}$$

Conversely, suppose that $\mathbb{E}\left[e^{\min\{2,|X_n-X|\}} - 1\right] \to 0$, take $\epsilon \in (0, 2)$, and estimate by the Chebychev (Markov) inequality

$$\begin{aligned}
\mathbb{P}\left(|X_n - X| \geq \epsilon\right) &= \mathbb{P}(\min\{2, |X_n - X|\} > \epsilon) \\
&= \mathbb{P}\left(e^{\min\{2,|X_n-X|\}} - 1 > e^{\epsilon} - 1\right) \\
&\leq \frac{1}{e^{\epsilon} - 1}\mathbb{E}\left[e^{\min\{2,|X-X_n|\}} - 1\right].
\end{aligned}$$

Here we have used that $e^{\min\{2,|X_n-X|\}} - 1 \geq 0$ almost surely.

By the condition of the problem,

$$\lim_{n\to\infty} \frac{1}{e^{\epsilon}-1}\mathbb{E}\left[e^{\min(2,|X-X_n|)}-1\right]=0,$$

and hence

$$\mathbb{P}(|X_n - X| > \epsilon) \to 0$$

as $n \to \infty$. This means that $X_n \to X$ in probability.

Probability Comprehensive Exam
January 15, 2016

Student Number: ☐

Instructions: Complete up to 5 of the 8 problems, and **circle** their numbers below – the uncircled problems will **not** be graded.

1 2 3 4 5 6 7 8

Write **only on the front side** of the solution pages. A **complete solution** of a problem is preferable to partial progress on several problems.

1. Let X have $\mathbb{E}X = 0$ and Var $X = \sigma^2 > 0$. Show that if $c > 0$, then

$$\mathbb{P}(X > c) \leq \frac{\sigma^2}{\sigma^2 + c^2}.$$

Hint: write $c - X = (c - X)_+ - (c - X)_-$ and use Cauchy-Schwarz type arguments.

Solution: Note that

$$c = \mathbb{E}(c - X) \leq \mathbb{E}(c - X)_+ = \mathbb{E}(c - X)\mathbf{1}_{\{X<c\}}.$$

Using Cauchy-Schwarz, the right side is bounded by

$$\sqrt{\mathbb{E}(c - X)^2\mathbb{P}(X < c)} = \sqrt{(c^2 + \sigma^2)\mathbb{P}(X < c)}.$$

Therefore

$$\mathbb{P}(X < c) \geq \frac{c^2}{c^2 + \sigma^2}$$

and the result follows.

2. Let $X = (X_1, X_2)$ be a Gaussian vector with zero mean and covariance matrix

$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix},$$

where $|\rho| < 1$. Find a matrix A such that $X = AZ$, where Z is a standard normal vector and derive the characteristic function of X as a function of ρ.

Solution: If $A : \mathbb{R}^2 \to \mathbb{R}^2$ is linear and we set $Y = AZ$, then we can compute the covariance matrix Σ of Y as

$$\Sigma_{i,j} = \mathbb{E}Y_iY_j = \sum_{k,l} \mathbb{E}A_{i,k}Z_kA_{j,l}Z_l = \sum_k A_{i,k}A_{j,k} = (AA^T)_{i,j}.$$

Therefore to find A in the statement of the problem, we need to solve $AA^T = \Sigma$. Putting $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, we want

$$\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} a^2 + b^2 & ac + bd \\ ac + bd & c^2 + d^2 \end{pmatrix}.$$

If we set $c = 0$ then this system becomes

$$\begin{aligned} a^2 + b^2 &= 1 \\ bd &= \rho \\ d^2 &= 1 \end{aligned}$$

If we set $d = 1, b = \rho$ and $a = \sqrt{1-\rho^2}$, then we obtain a solution.

To compute the characteristic function, we write

$$\mathbb{E}e^{it\cdot AZ} = \mathbb{E}e^{i\sum_j t_j(AZ)_j} = \mathbb{E}\exp\left(i\sum_{j,k} t_j A_{j,k} Z_k\right).$$

Now use independence of the Z_k's to obtain

$$\prod_k \mathbb{E}\exp\left(i\sum_j t_j A_{j,k} Z_k\right) = \prod_k \mathbb{E}\exp\left(i(A^T t)_k Z_k\right).$$

The inner term is the characteristic function of a standard normal random variable, evaluated at $(tA)_k$, so we obtain

$$\prod_k \exp\left(-\frac{1}{2}(A^T t)_k^2\right) = \exp\left(-\frac{1}{2}\|A^T t\|^2\right) = \exp\left(-\frac{1}{2}\langle AA^T t, t\rangle\right) = \exp\left(-\frac{1}{2}\langle \Sigma t, t\rangle\right).$$

In terms of ρ, this becomes

$$\phi_X(t_1, t_2) = \exp\left(-\frac{1}{2}(t_1^2 - 2\rho t_1 t_2 + t_2^2)\right).$$

3. Let $X_1, X_2, \ldots$ be i.i.d. uniform $(0,1)$ random variables. Show that

$$(X_1 \cdot \cdots \cdot X_n)^{1/n}$$

converges almost surely as $n \to \infty$ and compute the limit.

Solution: Let $Y_i = \log X_i$. Then if we write P_n for the above expression, one has

$$\log P_n = \frac{Y_1 + \cdots + Y_n}{n}.$$

If we can use the strong law of large numbers, then we will obtain

$$\log P_n \to \mathbb{E}Y_1, \text{ or } P_n \to e^{\mathbb{E}Y_1} \text{ almost surely.}$$

So we set to compute $\mathbb{E}Y_1$. Note that $Y_1 \le 0$ almost surely, so for $y \le 0$,

$$\mathbb{P}(Y_1 \le y) = \mathbb{E}(X_1 \le e^y) = e^y,$$

since X_1 is uniformly distributed. This means that $-Y_1$ is continuously distributed, nonnegative, and with $\mathbb{P}(-Y_1 \geq y) = \mathbb{P}(Y_1 \leq -y) = e^{-y}$ for $y \geq 0$. Using the tail-sum formula for expectation,

$$\mathbb{E}(-Y_1) = \int_0^\infty \mathbb{P}(-Y_1 \geq y)\ \mathrm{d}y = \int_0^\infty e^{-y}\ \mathrm{d}y = 1.$$

Therefore $\mathbb{E}Y_1 = -1$. We conclude that $\mathbb{E}|Y_1| = \mathbb{E}(-Y_1) = 1$ exists, so the strong law of large numbers applies and we obtain

$$(X_1 \cdot \cdots \cdot X_n)^{1/n} \to \exp\left(\mathbb{E}Y_1\right) = e^{-1}.$$

4. Let $X_1, X_2, \ldots$ be i.i.d. exponential variables with parameter 1 and set

$$M_n = \max\{X_1, \ldots, X_n\}.$$

Find sequences (a_n) and (b_n) of real numbers such that $(M_n - a_n)/b_n$ converges in distribution.

Solution: For $x \in \mathbb{R}$, by the i.i.d. assumption,

$$\mathbb{P}(M_n \leq x) = \mathbb{P}(X_i \leq x \text{ for all } i = 1, \ldots, n) = \mathbb{P}(X_1 \leq x)^n.$$

For $x \leq 0$, this is 0. For $x > 0$, one has

$$\mathbb{P}(X_1 \leq x) = \int_0^x e^{-t}\ \mathrm{d}t = 1 - e^{-x}.$$

Therefore

$$\mathbb{P}(M_n \leq x) = \begin{cases} 0 & \text{if } x \leq 0 \\ (1 - e^{-x})^n & \text{if } x > 0 \end{cases}.$$

Plugging in $x = \log n + c$, one obtains for n large

$$\mathbb{P}(M_n \leq \log n + c) = \left(1 - e^{-c-\log n}\right)^n = \left(1 - \frac{e^{-c}}{n}\right)^n \to e^{-e^{-c}}.$$

Note that $\lim_{c\to-\infty} e^{-e^{-c}} = 0$ and $\lim_{c\to\infty} e^{-e^{-c}} = 1$, and so since it is continuous (in particular right-continuous) the function $c \mapsto e^{-e^{-c}}$ is the distribution function for a probability measure. Therefore if we set Z to be a random variable with this distribution, and $a_n = \log n,\ b_n = 1$, one has

$$\mathbb{P}\left(\frac{M_n - a_n}{b_n} \leq c\right) = \mathbb{P}(M_n \leq \log n + c) \to e^{-e^{-c}},$$

and so $(M_n - a_n)/b_n \Rightarrow Z$.

5. Let $(N_t)_{t\geq 0}$ be a rate-λ Poisson process. Let $X_1, X_2, \ldots$ be i.i.d. random variables with $\mathbb{E}|X_1| < \infty$ (independent of the Poisson process as well) and define

$$S_t = \sum_{i=1}^{N_t} X_i.$$

Show that S_t/t converges in probability to a constant and compute this constant.

Solution: We can compute the characteristic function of S_t in terms of the characteristic function ϕ of X_1:

$$\begin{aligned}
\mathbb{E}e^{isS_t} = \sum_{n=0}^{\infty} \mathbb{E}e^{isS_t}\mathbf{1}_{\{N_t=n\}} &= \sum_{n=0}^{\infty} \mathbb{E}e^{is(X_1+\cdots+X_n)}\mathbf{1}_{\{N_t=n\}} \\
&= \sum_{n=0}^{\infty} \left(\mathbb{E}e^{isX_1}\right)^n \mathbb{P}(N_t = n) \\
&= e^{-\lambda t}\sum_{n=0}^{\infty} \phi^n(s)\frac{(\lambda t)^n}{n!} \\
&= e^{-\lambda t}\exp\left(\phi(s)\lambda t\right) \\
&= \exp\left(\lambda t(\phi(s)-1)\right).
\end{aligned}$$

So the characteristic function of S_t/t is

$$\exp\left(\lambda t(\phi(s/t)-1)\right) = \exp\left(\lambda s\left(\frac{\phi(s/t)-\phi(0)}{s/t}\right)\right).$$

As $t \to \infty$,

$$\frac{\phi(s/t)-\phi(0)}{s/t} \to \phi'(0) = \frac{\mathrm{d}}{\mathrm{d}a}\mathbb{E}e^{iaX_1} = i\mathbb{E}X_1.$$

So for each s,

$$\mathbb{E}e^{is(S_t/t)} \to e^{i\lambda s\mathbb{E}X_1},$$

which is the characteristic function of the constant variable $\lambda\mathbb{E}X_1$. By the continuity theorem, one has

$$S_t/t \Rightarrow \lambda\mathbb{E}X_1.$$

Since convergence in distribution to a constant implies convergence in probability, $S_t/t \to \lambda\mathbb{E}X_1$ in probability.

6. Let $X_1, X_2, \ldots$ be i.i.d. standard normal random variables and for $x \in (-1,1)$, set

$$Y = \sum_{n=1}^{\infty} x^n X_n.$$

Show that the sum defining Y converges and find its distribution.

Solution: To show that the sum converges, we can compute:

$$\mathbb{P}(|X_n| > n) = \frac{2}{\sqrt{2\pi}} \int_n^\infty e^{-t^2/2}\ \mathrm{d}t \leq Ce^{-n^2/2}.$$

(Here we are using the approximation $\mathbb{P}(X_n > x) \sim \frac{e^{-x^2/2}}{\sqrt{2\pi}x}$ and symmetry.) Therefore by Borel-Cantelli, since $\sum_n \mathbb{P}(|X_n| > n) < \infty$, one has

$$\mathbb{P}(|X_n| > n \text{ infinitely often}) = 0.$$

So we can dominate this sum by

$$\sum_{n=1}^\infty nx^n,$$

which converges for any $x \in (-1, 1)$. To find the limit, we compute the characteristic function. Let Y_n be the partial sum to term n and note that since $Y_n \to Y$ almost surely, also this convergence occurs in distribution, and therefore the characteristic function of Y_n converges pointwise to that of Y. Therefore

$$\phi_Y(t) = \mathbb{E}e^{itY} = \prod_{n=1}^\infty \mathbb{E}e^{itx^n X_n}.$$

Use the fact that the characteristic function for a standard Gaussian is $e^{-t^2/2}$ to obtain

$$\prod_{n=1}^\infty e^{-(tx^n)^2/2} = \exp\left(-(t^2/2)\sum_{n=1}^\infty x^{2n}\right) = \exp\left(-\frac{t^2 \frac{x^2}{1-x^2}}{2}\right).$$

This is the characteristic function of a Gaussian with mean zero and variance $x^2/(1-x^2)$.

7. Let $X_1, X_2, \ldots$ be independent random variables such that X_n has Binomial(n, p_n) distribution, for some $p_n > 0$. Show that if $np_n(1-p_n) \to \infty$, then

$$\frac{X_n - np_n}{\sqrt{np_n(1-p_n)}} \Rightarrow N(0, 1).$$

Solution: For $k \geq 1$, let $Y_{k,1}, \ldots, Y_{k,k}$ be i.i.d. Bernoulli random variables with parameter p_n. Then $Y_n = Y_{n,1} + \cdots + Y_{n,n}$ has the same distribution as X_n, has mean np_n, and has variance

$$\text{Var } X_n = \text{Var } (Y_{n,1} + \cdots + Y_{n,n}) = n\text{Var } Y_{n,1} = np_n(1 - p_n).$$

Thus the problem is asking us to show that

$$\frac{Y_n - \mathbb{E}Y_n}{\sqrt{\text{Var } Y_n}} \Rightarrow N(0, 1).$$

This will follow immediately once we show that hypothesis of Lindeberg's CLT hold. That is, we must show that if $s_n = \sqrt{\text{Var } Y_n}$, then for each $\epsilon > 0$,

$$\frac{1}{s_n^2} \sum_{i=1}^{n} \mathbb{E}Y_{n,i}^2 \mathbf{1}_{\{|Y_{n,i}| \geq \epsilon s_n\}} \to 0 \text{ as } n \to \infty.$$

As $np_n(1 - p_n) \to \infty$, one has $s_n \to \infty$. So for all large n, $\epsilon s_n > 1$. Since the $Y_{n,i}$'s are bounded in absolute value by 1, the indicator function is 0 for all large n. Thus the above expression is 0 for all large n and we are done.

8. A sequence of events $A_1, A_2, \ldots$ is said to be 1-dependent if for every $k \geq 1$, the sigma-algebras $\sigma(A_1, \ldots, A_k)$ and $\sigma(A_{k+2}, A_{k+3}, \ldots)$ are independent. Prove that if $A_1, A_2, \ldots$ are 1-dependent and E is a tail event:

$$E \in \cap_n \sigma(A_n, A_{n+1}, \ldots),$$

then $\mathbb{P}(E) = 0$ or 1.

Solution: Here the proof is almost exactly the same as that of Kolmogorov's 0/1 law. So we will use some of the tools from that proof. We aim to show that E is independent of itself, so $\mathbb{P}(E) = \mathbb{P}(E)^2$ and the result will follow.

As in the proof of the Kolmogorov 0/1 law, if we define the collection

$$\mathcal{C}_E = \{A : \mathbb{P}(A \cap E) = \mathbb{P}(A)\mathbb{P}(E)\},$$

then $\mathcal{C}_E$ is a λ-system. We claim that it contains the π-system

$$\Pi = \cup_n \sigma(A_1, A_2, \ldots, A_n).$$

Indeed, if $A \in \Pi$, then $A \in \sigma(A_1, \ldots, A_n)$ for some n. Since E is a tail event,

$$E \in \cap_k \sigma(A_k, A_{k+1}, \ldots) \subset \sigma(A_{n+2}, A_{n+3}, \ldots).$$

By assumption, $\sigma(A_1, A_2, \ldots, A_n)$ is independent of $\sigma(A_{n+2}, \ldots)$, and so E is independent of A, giving $\Pi \subset \mathcal{C}_E$.

By the $\pi - \lambda$ theorem, one has $\mathcal{C}_E \supset \sigma(\Pi)$. However

$$\sigma(\Pi) = \sigma(\cup_n \sigma(A_1, \ldots, A_n)) = \sigma(A_1, A_2, \ldots),$$

and this last sigma-algebra contains E. Therefore $E \in \mathcal{C}_E$ and E is independent of itself.

Probability Comprehensive Exam
January 18, 2017

Student Number:

Instructions: Complete 5 of the 8 problems, and **circle** their numbers below – the uncircled problems will **not** be graded.

1 2 3 4 5 6 7 8

Write **only on the front side** of the solution pages. A **complete solution** of a problem is preferable to partial progress on several problems.

1. Show that if X_n and Y_n are independent for $n = 1, 2, ...$ and $X_n \to^w X$, $Y_n \to^w Y$, where X and Y are independent, then $X_n + Y_n \to^w X + Y$.

Solution: Since X and Y are independent, the characteristic function of $X + Y$ evaluated at $t \in \mathbb{R}$ is

$$\phi_{X+Y}(t) = \mathbb{E}e^{it(X+Y)} = \mathbb{E}e^{itX}e^{itY} = \mathbb{E}e^{itX}\mathbb{E}e^{itY} = \phi_X(t)\phi_Y(t).$$

On the other hand, since $X_n \to^w X$ and $Y_n \to^w Y$, one has $\phi_{X_n}(t) \to \phi_X(t)$ and $\phi_{Y_n}(t) \to \phi_Y(t)$ for each t. Again using independence of X_n and Y_n,

$$\phi_{X_n+Y_n}(t) = \phi_{X_n}(t)\phi_{Y_n}(t) \to \phi_X(t)\phi_Y(t) = \phi_{X+Y}(t).$$

Since the characteristic function of $X_n + Y_n$ converges pointwise to that of $X + Y$, we conclude that $X_n + Y_n \to^w X + Y$.

2. Let X be a random variable with mean zero and finite variance σ^2. Prove that for every $c > 0$,

$$P(X > c) \leq \frac{\sigma^2}{\sigma^2 + c^2}.$$

Hint: Combine the inequality $\mathbb{E}(c - X) \leq \mathbb{E}\left((c - X)\mathbf{1}_{\{X<c\}}\right)$ with the Cauchy-Schwartz inequality.

Solution: By Cauchy-Schwarz,

$$c = \mathbb{E}(c - X) \leq \mathbb{E}(c - X)\mathbf{1}_{\{X<c\}} \leq \sqrt{\mathbb{E}(c - X)^2\mathbb{P}(X < c)}.$$

However

$$\mathbb{E}(c - X)^2 = \mathbb{E}(c^2 - 2cX + X^2) = c^2 + \sigma^2,$$

so

$$c \leq \sqrt{(c^2 + \sigma^2)\mathbb{P}(X < c)},$$

or

$$\mathbb{P}(X < c) \geq \frac{c^2}{c^2 + \sigma^2}.$$

3. Let $X_1, X_2, ...$ be i.i.d. random variables uniformly distributed on $[0, 1]$. Show that with probability 1,

$$\lim_{n\to\infty} (X_1 \cdot \cdots \cdot X_n)^{\frac{1}{n}}$$

exists and compute its value.

Solution: Define $Y_n = \log X_n$, so that the quantity we are considering is

$$\lim_{n\to\infty} \exp\left(\frac{1}{n}\sum_{i=1}^{n} Y_i\right).$$

We can compute $\mathbb{E}Y_n$ as

$$\mathbb{E}Y_n = \int_0^1 \log x \,\mathrm{d}x = -1,$$

so since (Y_n) is an i.i.d. sequence with entries of mean -1, the strong law of large numbers gives $\frac{1}{n}\sum_{i=1}^{n} Y_i \to -1$ a.s. Since $x \mapsto e^x$ is continuous, we obtain a.s.

$$\lim_{n\to\infty} (X_1 \cdot \dots \cdot X_n)^{1/n} = \exp\left(\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} Y_i\right) = e^{-1}.$$

4. Let X and Y be independent and suppose that each has a uniform distribution on $(0,1)$. Let $Z = \min\{X, Y\}$. Find the density $f_Z(z)$ for Z.

Solution: Let $z \in (0,1)$. By independence,

$$\mathbb{P}(Z > z) = \mathbb{P}(X > z \text{ and } Y > z) = \mathbb{P}(X > z)\mathbb{P}(Y > z) = \mathbb{P}(X > z)^2 = (1-z)^2.$$

Therefore the distribution function of Z is $F(z) = \mathbb{P}(Z \le z) = 1 - (1-z)^2$ when $z \in (0,1)$. It is easy to see that $F(z) = 0$ if $z \le 0$ and $F(z) = 1$ if $z \ge 1$. To compute the density, we take the derivative:

$$f_Z(z) = \frac{\mathrm{d}}{\mathrm{d}z} F(z) = 2(1-z),$$

whenever $z \in (0,1)$, and zero otherwise.

5. Show that the characteristic function φ of a random variable X is real if and only if X and $-X$ have the same distribution.

Solution: The characteristic function ϕ of $-X$ evaluated at $t \in \mathbb{R}$ is

$$\phi(t) = \mathbb{E}e^{it(-X)} = \mathbb{E}e^{-itX} = \overline{\mathbb{E}e^{itX}} = \overline{\varphi(t)}.$$

Here we have used that for a complex variable $U + iW$, one has

$$\overline{\mathbb{E}(U+iW)} = \overline{\mathbb{E}U + i\mathbb{E}W} = \mathbb{E}U - i\mathbb{E}W = \mathbb{E}(U - iW) = \mathbb{E}(\overline{U+iW}).$$

If X and $-X$ have the same distribution, then their characteristic functions are equal, so $\varphi(t) = \overline{\varphi(t)}$ for all t, meaning φ is real. Conversely, if φ is real, then $\varphi(t) = \overline{\varphi(t)}$ for all t, meaning $\phi = \varphi$. Since the characteristic functions of X and $-X$ then are equal, the variables have the same distribution.

6. Let X_i be i.i.d. random variables uniformly distributed on $[0, 2]$. Let $S_n = X_1 + \cdots + X_n$. Show that
$$\frac{3\sqrt{3}}{2} n^{\frac{1}{6}} \left(\sqrt[3]{S_n} - \sqrt[3]{n}\right) \to^w Z,$$
where Z is a standard normal random variable.

Solution: First, observe that $\mathbb{E}X_i = 1$ and $\sigma := \sqrt{\text{Var } X_i} = \frac{2}{\sqrt{3}}$. Therefore, by the CLT, the random variable $\frac{S_n - n}{\frac{2}{\sqrt{3}}\sqrt{n}}$ converges weakly to a standard Gaussian random variable.

We estimate the probability

$$\begin{aligned}
&P\left(\frac{3\sqrt{3}}{2} n^{\frac{1}{6}} \left(\sqrt[3]{S_n} - \sqrt[3]{n}\right) \le t\right) \\
&= P\left(\sqrt[3]{S_n} \le \frac{2}{3\sqrt{3}n^{\frac{1}{6}}} t + \sqrt[3]{n}\right) \\
&= P\left(\frac{S_n - n}{\frac{2}{\sqrt{3}}\sqrt{n}} \le t + O(\frac{1}{\sqrt{n}})\right) \\
&= P\left(\frac{S_n - n}{\frac{2}{\sqrt{3}}\sqrt{n}} \le t\right) + P\left(\frac{S_n - n}{\frac{2}{\sqrt{3}}\sqrt{n}} \in (t, t + O(\frac{1}{\sqrt{n}}))\right).
\end{aligned}$$

The second summand tends to zero as $n \to \infty$: indeed, for every $\epsilon > 0$ there exists an n large enough so that $O(\frac{1}{\sqrt{n}}) < \epsilon$, and hence

$$P\left(\frac{S_n - n}{\frac{2}{\sqrt{3}}\sqrt{n}} \in (t, t + O(\frac{1}{\sqrt{n}}))\right) \le P\left(\frac{S_n - n}{\frac{2}{\sqrt{3}}\sqrt{n}} \in (t, t + \epsilon)\right) \to \frac{1}{\sqrt{2\pi}} \int_t^{t+\epsilon} e^{-\frac{t^2}{2}} dt,$$

which can be made arbitrarily small by choosing small enough ϵ.

The first summand tends to $P(Z \le t)$, and hence

$$P\left(\frac{3\sqrt{3}}{2} n^{\frac{1}{6}} \left(\sqrt[3]{S_n} - \sqrt[3]{n}\right) \le t\right) \to_{n\to\infty} P(Z \le t),$$

which implies weak convergence, since the distribution of Z is continuous.

7. Let $v = \left(\frac{1}{\sqrt{n}}, \ldots, \frac{1}{\sqrt{n}}\right)$ be a unit vector in $\mathbf{R}^n$. Consider the set A in $\mathbf{R}^n$ be given by

$$A = \left\{ x \in \mathbf{R}^n : \, x_i \in \left[-\frac{1}{2}, \frac{1}{2}\right], \langle x, v\rangle \le \frac{t}{2\sqrt{3}} \right\}.$$

Prove that as the dimension $n \to \infty$,

$$Vol_n(A) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{x^2}{2}} dx + O(\frac{1}{\sqrt{n}}).$$

Solution: Consider a random vector X uniformly distributed on $[-\frac{1}{2}, \frac{1}{2}]^n$. Its coordinates are i.i.d., with $\mathbb{E}X_i = 0$, $\sqrt{\text{Var } X_i} := \sigma = \frac{1}{2\sqrt{3}}$ and $\mathbb{E}|X_i|^3 = \frac{1}{32} < +\infty$. Therefore, by Berry-Essen's theorem,

$$\left| P\left(\frac{X_1 + \ldots + X_n}{\sigma\sqrt{n}} \le t\right) - P\left(Z \le t\right) \right| \le O(\frac{1}{\sqrt{n}}).$$

It remains to observe that

$$Vol_n(A) = P\left(\langle X, v\rangle \le \frac{t}{2\sqrt{3}}\right) = P\left(\frac{X_1 + \ldots + X_n}{\sigma\sqrt{n}} \le t\right).$$

8. Assume $X_1, X_2, \ldots, X_n, \ldots$ are i.i.d. standard normal random variables. Show without using the law of the iterated logarithm that for any $\lambda > 1/2$,

$$\frac{1}{n^{\lambda}}(X_1 + \ldots + X_n) \to^{a.s.} 0$$

Solution: The sum of n standard normal variables is normal with mean zero and variance n, as can be seen from computing characteristic functions: the characteristic function of the sum is, by independence,

$$\mathbb{E}e^{it(X_1+\cdots+X_n)} = \left(\mathbb{E}e^{itX_1}\right)^n = \left(e^{-t^2/2}\right)^n = e^{-t^2n/2},$$

which is the characteristic function of a Gaussian with mean zero and variance n. So we can compute for $\epsilon > 0$

$$\mathbb{P}\left(\left|\frac{1}{n^\lambda}(X_1+\cdots+X_n)\right| > \epsilon\right) = \mathbb{P}(|Z_n| > \epsilon n^\lambda),$$

where Z_n is Gaussian with mean zero and variance n. If Z is a standard normal variable, then $\sqrt{n}Z$ has the same distribution as Z_n, so this probability is

$$\mathbb{P}(|Z| > \epsilon n^{\lambda-1/2}),$$

or for $\sigma = 1/(\lambda - 1/2) > 0$,

$$\mathbb{P}\left(\frac{|Z|^\sigma}{\epsilon^\sigma} > n\right).$$

However a standard Gaussian has finite moments of all orders, so we use the characterization for a nonnegative random variable Y of $\mathbb{E}Y < \infty \Leftrightarrow \sum_n \mathbb{P}(Y > n) < \infty$ to say that since $\mathbb{E}\frac{|Z|^\sigma}{\epsilon^\sigma} < \infty$, one has

$$\sum_n \mathbb{P}\left(\frac{|Z|^\sigma}{\epsilon^\sigma} > n\right) < \infty.$$

This implies

$$\sum_n \mathbb{P}\left(\left|\frac{1}{n^\lambda}(X_1+\cdots+X_n)\right| > \epsilon\right) < \infty,$$

and so by Borel-Cantelli, a.s. $\left|\frac{1}{n^\lambda}(X_1+\cdots+X_n)\right| > \epsilon$ for only finitely many n. This implies convergence to 0 a.s.

Probability Comprehensive Exam
Fall 2018

Student Number: ☐

Instructions: Complete 5 of the 9 problems, and **circle** their numbers below – the uncircled problems will **not** be graded.

1 2 3 4 5 6 7 8 9

Write **only on the front side** of the solution pages. A **complete solution** of a problem is preferable to partial progress on several problems.

1. Use the SLLN to find the following limit:

$$\lim_{n\to\infty}\int_0^1\cdots\int_0^1 \frac{x_1^2+\cdots+x_n^2}{x_1+\cdots+x_n}dx_1\ldots dx_n.$$

Solution: Let $U_1,\ldots,U_n$ be i.i.d. random variables with uniform distribution in $[0,1]$. Then

$$\mathbb{E}\frac{U_1^2+\cdots+U_n^2}{U_1+\cdots+U_n}=\int_0^1\cdots\int_0^1 \frac{x_1^2+\cdots+x_n^2}{x_1+\cdots+x_n}dx_1\ldots dx_n.$$

By the SLLN,

$$\frac{U_1^2+\cdots+U_n^2}{n}\to\mathbb{E}U_1^2=\int_0^1 x^2dx=\frac{1}{3}\text{ a.s.}$$

and

$$\frac{U_1+\cdots+U_n}{n}\to\mathbb{E}U_1=\int_0^1 xdx=\frac{1}{2}\text{ a.s.}$$

Therefore,

$$\frac{U_1^2+\cdots+U_n^2}{U_1+\cdots+U_n}=\frac{(U_1^2+\cdots+U_n^2)/n}{(U_1+\cdots+U_n)/n}\to\frac{2}{3}\text{ a.s.}$$

Since also

$$0\le\frac{U_1^2+\cdots+U_n^2}{U_1+\cdots+U_n}\le 1,$$

we have, by Lebesgue dominated convergence, that

$$\int_0^1\cdots\int_0^1 \frac{x_1^2+\cdots+x_n^2}{x_1+\cdots+x_n}dx_1\ldots dx_n=\mathbb{E}\frac{U_1^2+\cdots+U_n^2}{U_1+\cdots+U_n}\to\frac{2}{3}$$

as $n\to\infty$.

2. Suppose $X_1,\ldots,X_n$ are i.i.d. random variables such that $\mathbb{P}\{X_j=+1\}=\mathbb{P}\{X_j=-1\}=1/2$. Let $S_k:=X_1+\cdots+X_k, k=1,\ldots,n$. Prove that

$$\mathbb{P}\{\max_{1\le k\le n}S_k\ge l\}=2\mathbb{P}\{S_n>l\}+\mathbb{P}\{S_n=l\}.$$

Solution: Note that, by additivity and independence,

$$\begin{aligned}\mathbb{P}\{S_n > l\} &= \sum_{k=1}^{n} \mathbb{P}\{S_1 < l, \ldots, S_{k-1} < l, S_k = l, S_n - S_k > 0\} \\ &= \sum_{k=1}^{n} \mathbb{P}\{S_1 < l, \ldots, S_{k-1} < l, S_k = l\}\mathbb{P}\{S_n - S_k > 0\}\end{aligned}$$

and

$$\begin{aligned}\mathbb{P}\{S_n = l\} &= \sum_{k=1}^{n} \mathbb{P}\{S_1 < l, \ldots, S_{k-1} < l, S_k = l, S_n - S_k = 0\} \\ &= \sum_{k=1}^{n} \mathbb{P}\{S_1 < l, \ldots, S_{k-1} < l, S_k = l\}\mathbb{P}\{S_n - S_k = 0\}.\end{aligned}$$

This implies that

$$\begin{aligned}&2\mathbb{P}\{S_n > l\} + \mathbb{P}\{S_n = l\} \\ &= \sum_{k=1}^{n} \mathbb{P}\{S_1 < l, \ldots, S_{k-1} < l, S_k = l\}(2\mathbb{P}\{S_n - S_k > 0\} + \mathbb{P}\{S_n - S_k = 0\}).\end{aligned}$$

Finally, note that by symmetry of r.v. $S_n - S_k = \sum_{j=k+1}^{n} X_j$,

$$\begin{aligned}&2\mathbb{P}\{S_n - S_k > 0\} + \mathbb{P}\{S_n - S_k = 0\} \\ &= \mathbb{P}\{S_n - S_k > 0\} + \mathbb{P}\{S_n - S_k < 0\} + \mathbb{P}\{S_n - S_k = 0\} = 1\end{aligned}$$

and

$$\sum_{k=1}^{n} \mathbb{P}\{S_1 < l, \ldots, S_{k-1} < l, S_k = l\} = \mathbb{P}\{\max_{1\le k\le n} S_k \ge l\},$$

implying the claim.

3. Let $\{Z_n\}$ be i.i.d. standard normal r.v. and let $\{a_n\}$ be a sequence of nonnegative real numbers. Prove that $\sum_{n=1}^{\infty} a_n Z_n^2 < +\infty$ a.s. if and only if $\sum_{n=1}^{\infty} a_n < +\infty$.

Solution: If $\sum_{n=1}^{\infty} a_n < +\infty$, then

$$\mathbb{E}\sum_{n=1}^{\infty} a_n Z_n^2 = \sum_{n=1}^{\infty} a_n \mathbb{E} Z_n^2 = \sum_{n=1}^{\infty} a_n < +\infty,$$

implying that the nonnegative r.v. $\xi := \sum_{n=1}^{\infty} a_n Z_n^2$ is finite a.s. On the other hand, if $\xi < +\infty$ a.s., then $e^{-\xi} > 0$ a.s., implying that $\mathbb{E}e^{-\xi} > 0$. By a straightforward computation,

$$\mathbb{E}e^{-\xi} = \prod_{n=1}^{\infty} \mathbb{E}e^{-a_n Z_n^2} = \prod_{n=1}^{\infty} \mathbb{E}e^{-a_n Z_1^2} = \prod_{n=1}^{\infty} \frac{1}{\sqrt{1+2a_n}}$$

The last product is strictly positive if and only if the series $\sum_{n=1}^{\infty} \log(1+2a_n)$ converges, which implies $\sum_{n=1}^{\infty} a_n < +\infty$.

4. Let φ be the characteristic function of r.v. X. Show that

$$\psi_1(t) = |\varphi(t)|^2 \text{ and } \psi_2(t) = \frac{1}{t}\int_0^t \varphi(s)ds$$

are also characteristic functions.

Solution: Note that

$$\psi_1(t) = \varphi(t)\overline{\varphi(t)} = \mathbb{E}e^{itX}\mathbb{E}e^{-itX} = \mathbb{E}e^{itX}\mathbb{E}e^{-itY} = \mathbb{E}e^{it(X-Y)},$$

where Y is an independent copy of X. Thus, ψ_1 is the characteristic function of $X - Y$.

By change of variable and the properties of conditional expectation,

$$\psi_2(t) = \frac{1}{t}\int_0^t \varphi(s)ds = \int_0^1 \varphi(tu)du = \int_0^1 \mathbb{E}e^{itXu}du$$
$$= \int_0^1 \mathbb{E}(e^{itXU}|U=u)du = \mathbb{E}\mathbb{E}(e^{itXU}|U) = \mathbb{E}e^{itXU},$$

where U is a random variable with uniform distribution in $[0,1]$ independent of X. Thus, ψ_2 is the characteristic function of XU.

5. For distribution functions F, G on the real line, define

$$L(F,G) := \inf\Big\{\varepsilon > 0 : \forall t \in \mathbb{R}\ F(t) \le G(t+\varepsilon)+\varepsilon, G(t) \le F(t+\varepsilon)+\varepsilon\Big\}.$$

It is known that L is a metric. Prove that $L(F_n, F) \to 0$ as $n \to \infty$ if and only if F_n converges weakly to F.

Solution: If $L(F_n, F) \to 0$ as $n \to \infty$, then, for any $\varepsilon > 0$ and all large enough n, $L(F_n, F) < \varepsilon$. This implies that, for all large enough n,

$$\forall t \; F(t-\varepsilon) - \varepsilon \le F_n(t) \le F(t+\varepsilon) + \varepsilon.$$

Therefore

$$F(t-\varepsilon) - \varepsilon \le \liminf_{n\to\infty} F_n(t) \le \limsup_{n\to\infty} F_n(t) \le F(t+\varepsilon) + \varepsilon. \tag{1}$$

Passing to the limit when $\varepsilon \to 0$, we get

$$F(t-) \le \liminf_{n\to\infty} F_n(t) \le \limsup_{n\to\infty} F_n(t) \le F(t). \tag{2}$$

If t is a continuity point of F, we have $F(t) = F(t-)$ and

$$\lim_{n\to\infty} F_n(t) = F(t),$$

which implies the weak convergence of F_n to F.

On the other hand, the weak convergence of F_n to F easily implies (2), which implies (1). It follows from (1) and the definition of L that $L(F_n, F) < 2\varepsilon$ for all n large enough. Therefore, $L(F_n, F) \to 0$ as $n \to \infty$.

6. Let $X_1, X_2, \ldots, X_n, \ldots$ be identically distributed (not necessarily independent!) random variables with finite first moment. Is the following,

$$n^{-1}\mathbb{E} \max_{1\le k\le n} |X_k| \longrightarrow 0,$$

as $n \to +\infty$, true or false?

Solution: True! Indeed, for any $A > 0$, and using the identical distribution assumption,

$$
\begin{aligned}
\mathbb{E}\max_{1\le k\le n}|X_k| &= \int_0^{+\infty}\mathbb{P}(\max_{1\le k\le n}|X_k|>t)dt\\
&= \int_0^{A}\mathbb{P}(\max_{1\le k\le n}|X_k|>t)dt + \int_A^{+\infty}\mathbb{P}(\max_{1\le k\le n}|X_k|>t)dt\\
&\le A + \int_A^{+\infty}\sum_{k=1}^{n}\mathbb{P}(|X_k|>t)dt\\
&= A + n\int_A^{+\infty}\mathbb{P}(|X_1|>t)dt.
\end{aligned}
$$

Therefore, for any $A>0$,

$$
\limsup_{n\to+\infty}\frac{1}{n}\mathbb{E}\max_{1\le k\le n}|X_k| \le \int_A^{+\infty}\mathbb{P}(|X_1|>t)dt.
$$

But, $\mathbb{E}|X_1| = \int_0^{+\infty}\mathbb{P}(|X_1|>t)dt < +\infty$, and so by dominated convergence,

$$
\limsup_{A\to+\infty}\limsup_{n\to+\infty}\frac{1}{n}\mathbb{E}\max_{1\le k\le n}|X_k| \le 0,
$$

which gives the result.

7. Let $X_1, X_2, \ldots, X_n, \ldots$ be iid random variables with common characteristic function φ and let $S_n = \sum_{k=1}^n X_k$. Show that if φ is differentiable at 0 with $\varphi'(0) = i\mu$, then, as $n \to +\infty$, $S_n/n \to \mu$, in probability.

Solution: In case the limit is degenerate then convergence in probability is equivalent to weak convergence. In other words, $S_n/n \to \mu$, in probability if and only if $S_n/n \Rightarrow \mu$. In turn by the Lévy continuity theorem, this last condition is equivalent to the requirement that for all $t\in\mathbb{R}$, $\mathbb{E}(e^{itS_n/n}) \to e^{it\mu}$. Now by the iid assumption, $\mathbb{E}(e^{itS_n/n}) = (\varphi(t/n))^n$. Since φ is differentiable at 0,

$$
\lim_{n\to+\infty}\frac{\varphi(t/n)-1}{t/n} = \varphi'(0) = i\mu,
$$

i.e., $\lim_{n\to+\infty} n(\varphi(t/n)-1) = it\mu$. Finally, since

$$
(\varphi(t/n))^n = \left(1+\frac{n(\varphi(t/n)-1)}{n}\right)^n,
$$

using complex logarithms or the fact that if $z_n\in\mathbb{C}$ is such that $z_n\to z\in\mathbb{C}$, then $(1+z_n/n)^n \to e^z$, the result follows.

8. Let X and Y be two independent and positive random variables with respective density f_X and f_Y and let $g : (0, +\infty) \longrightarrow (0, +\infty)$, be a bounded Borel function. Find

$$\mathbb{E}\left(g\left(\frac{X}{Y}\right)|Y\right),$$

the conditional expectation of $g(X/Y)$ given Y and then infer that $V = X/Y$ has a density that you will identify.

Solution: Since X and Y are independent, $\mathbb{E}\left(g\left(\frac{X}{Y}\right)|Y\right) = h(Y)$, with $h(y) = \mathbb{E}\left(g(X/y)\right)$. Therefore,

$$\begin{aligned} h(y) &= \int_0^{+\infty} g\left(\frac{x}{y}\right) f_X(x)dx \\ &= y\int_0^{+\infty} g(v)f_X(yv)dv. \end{aligned}$$

Next, for any g as above,

$$\mathbb{E}g(V) = \mathbb{E}(\mathbb{E}(g(V)|Y)) = \mathbb{E}h(Y).$$

But, using the Fubini-Tonelli Theorem which is valid since all our functions are non-negative as well as Lebesgue measurable,

$$\begin{aligned} \mathbb{E}g(V) = \mathbb{E}h(Y) &= \int_0^{+\infty} h(y)f_Y(y)dy \\ &= \int_0^{+\infty} f_Y(y)\left(\int_0^{+\infty} yg(v)f_X(yv)dv\right)dy \\ &= \int_0^{+\infty} g(v)\left(\int_0^{+\infty} yf_Y(y)f_X(yv)dy\right)dv \\ &= \int_0^{+\infty} g(v)f(v)dv, \end{aligned}$$

where $f(v) := \int_0^{+\infty} yf_Y(y)f_X(yv)dy$ is therefore the density of V.

9. Let X, Y, Z be random variables such that (X, Z) and (Y, Z) are identically distributed. Let $f : \mathbb{R} \longrightarrow \mathbb{R}$ be a Borel function such that $f(X)$ is integrable.
(i) Show that $\mathbb{E}(f(X)|Z) = \mathbb{E}(f(Y)|Z), a.s.$
(ii) Let $T_1, T_2, \ldots T_n$ be iid random variables with finite first moment and let $T = T_1 + \cdots + T_n$. Using (i) show that

$$\mathbb{E}(T_1|T) = \frac{T}{n}.$$

Solution: (i) For any non-negative (or bounded) Borel function g, since $g(Z)$ is Z-measurable, since the "expectation of the conditional expectation is the expectation", and using the identical distribution assumption,

$$\begin{aligned}\mathbb{E}(g(Z)\mathbb{E}(f(X)|Z)) &= \mathbb{E}(\mathbb{E}(g(Z)f(X)|Z)) = \mathbb{E}(g(Z)f(X)) \\ &= \mathbb{E}(g(Z)f(Y)) = \mathbb{E}(\mathbb{E}(g(Z)f(Y)|Z)) \\ &= \mathbb{E}(g(Z)\mathbb{E}(f(Y)|Z)),\end{aligned}$$

from which it follows (by the very definition and uniqueness of the conditional expectation) that $\mathbb{E}(f(X)|Z) = \mathbb{E}(f(Y)|Z), a.s.$, since both quantities above are $Z-$measurable.

(ii) Clearly, $(T_1, T), (T_2, T), \ldots, (T_n, T)$ are identically distributed and so, by (i),

$$\mathbb{E}(T_1|T) = \mathbb{E}(T_2|T) = \cdots = \mathbb{E}(T_n|T).$$

Therefore,

$$\begin{aligned}n\mathbb{E}(T_1|T) &= \mathbb{E}(T_1|T) + \mathbb{E}(T_2|T) + \cdots + \mathbb{E}(T_n|T) \\ &= \mathbb{E}(T_1 + T_2 + \cdots + T_n|T) \\ &= \mathbb{E}(T|T) = T,\end{aligned}$$

which shows that $\mathbb{E}(T_1|T) = T/n$.

Probability Comprehensive Exam
Spring 2018

Student Number: ☐

Instructions: Complete 5 of the 9 problems, and **circle** their numbers below – the uncircled problems will **not** be graded.

1 2 3 4 5 6 7 8 9

Write **only on the front side** of the solution pages. A **complete solution** of a problem is preferable to partial progress on several problems.

1. Let $\{X_n\}$ be a sequence of independent identically distributed random variables with exponential distribution (in other words, $X_n \geq 0$ a.s. and $\mathbb{P}\{X_n \geq t\} = e^{-\lambda t}, t \geq 0$ for some $\lambda > 0$). Prove that

$$\limsup_{n\to\infty} \frac{X_n}{\log n} < \infty \text{ a.s.}$$

Solution: For $C > 0$, $\mathbb{P}\{X_n \geq C \log n\} = e^{-C\lambda \log n} = n^{-C\lambda}$. For $C > \lambda^{-1}$,

$$\sum_{n\geq 1} \mathbb{P}\{X_n \geq C \log n\} = \sum_{n\geq 1} n^{-C\lambda} < \infty.$$

By Borel-Cantelli Lemma,

$$\mathbb{P}\{X_n \geq C \log n \text{ infinitely often}\} = 0,$$

implying the claim.

2. Suppose f is a continuous function on $[0,1]$. Use the Law of Large Numbers to prove that

$$\lim_{n\to\infty} \int_0^1 \cdots \int_0^1 f((x_1 \dots x_n)^{1/n}) dx_1 \dots dx_n = f\left(\frac{1}{e}\right).$$

Solution: Let $X_1, \dots, X_n$ be i.i.d. random variables with uniform distribution in $[0,1]$. Then

$$\int_0^1 \cdots \int_0^1 f((x_1 \dots x_n)^{1/n}) dx_1 \dots dx_n = \mathbb{E} f((X_1 \dots X_n)^{1/n})$$
$$= \mathbb{E} f\left(\exp\left\{\frac{\log X_1 + \cdots + \log X_n}{n}\right\}\right).$$

By the Strong Law of Large Numbers,

$$\frac{\log X_1 + \cdots + \log X_n}{n} \to \mathbb{E} \log X_1 = \int_0^1 \log x dx = -1 \text{ as } n \to \infty \text{ a.s.}$$

By continuity of f and Lebesgue dominated convergence theorem,

$$\mathbb{E} f\left(\exp\left\{\frac{\log X_1 + \cdots + \log X_n}{n}\right\}\right) \to f(\exp\{-1\}),$$

implying the result.

3. Let X, Y be random variables with $\mathbb{E}|X| < \infty, \mathbb{E}|Y| < \infty$. If $\mathbb{E}(X|Y) = Y$ and $\mathbb{E}(Y|X) = X$ a.s., then $X = Y$ a.s. Prove it.

Solution: Let f be a uniformly bounded strictly increasing function. It follows from the assumptions that

$$\mathbb{E}(X-Y)f(Y) = \mathbb{E}\mathbb{E}(X-Y|Y)f(Y) = \mathbb{E}(\mathbb{E}(X|Y)-Y)f(Y) = \mathbb{E}(Y-Y)f(Y) = 0.$$

Similarly, $\mathbb{E}(X-Y)f(X) = 0$, which implies

$$\mathbb{E}(X-Y)(f(X)-f(Y)) = 0.$$

Since f is strictly increasing, $(X-Y)(f(X)-f(Y)) \geq 0$ and, moreover, $(X-Y)(f(X)-f(Y)) = 0$ if and only if $X = Y$. Therefore, we have $(X-Y)(f(X)-f(Y)) = 0$ a.s., implying $X = Y$ a.s.

4. Let $X_1, \ldots, X_n$ be i.i.d. random variables with mean μ and variance $\sigma^2 < +\infty$. Let f be a function continuously differentiable at the point μ. Prove that the sequence of random variables

$$n^{1/2}\left(f\left(\frac{X_1 + \cdots + X_n}{n}\right) - f(\mu)\right)$$

converges in distribution to a normal random variable. What is the mean and the variance of the limit?

Solution: Let $Y_n = n^{1/2}(\frac{X_1+\cdots+X_n}{n} - \mu)$. By the Central Limit Theorem, Y_n converges in distribution to a normal random variable Y with mean zero and variance σ^2 as $n \to \infty$ and, by the Law of Large Numbers, $n^{-1/2}Y_n \to 0$ as $n \to \infty$ in probability. By the first order Taylor expansion,

$$f(x) = f(\mu) + f'(\mu)(x-\mu) + r(\mu; x-\mu)(x-\mu),$$

where $r(\mu;\delta) \to 0$ as $\delta \to 0$. Therefore,

$$\begin{aligned} n^{1/2}\left(f\left(\frac{X_1 + \cdots + X_n}{n}\right) - f(\mu)\right) &= n^{1/2}(f(\mu + n^{-1/2}Y_n) - f(\mu)) \\ &= f'(\mu)Y_n + r(\mu; n^{-1/2}Y_n)Y_n. \end{aligned}$$

Since $f'(\mu)Y_n$ converges in distribution to $f'(\mu)Y$ and $r(\mu; n^{-1/2}Y_n)Y_n$ converges in probability to 0, we can conclude that $n^{1/2}\left(f\left(\frac{X_1+\cdots+X_n}{n}\right) - f(\mu)\right)$ converges in distribution to a normal random variable with mean 0 and variance $(f'(\mu))^2\sigma^2$.

5. Let $X_1, \ldots, X_n, \ldots$ be i.i.d. random variables with $\mathbb{E}X_1 = 0$ and $\mathrm{Var}(X_1) = 1$. Let $S_n = X_1 + \cdots + X_n$. Prove that

$$\limsup_{n\to\infty} \frac{S_n}{\sqrt{n}} = +\infty.$$

Solution: By the Central Limit Theorem, for all $A > 0$,

$$\lim_{N\to\infty} \mathbb{P}\left\{ \frac{S_N}{\sqrt{N}} \geq A \right\} = \mathbb{P}\{Z \geq A\} > 0,$$

where Z is a standard normal random variable. Therefore,

$$\mathbb{P}\left\{ \limsup_{n\to\infty} \frac{S_n}{\sqrt{n}} \geq A \right\} = \mathbb{P}\left\{ \lim_{N\to\infty} \sup_{n\geq N} \frac{S_n}{\sqrt{n}} \geq A \right\}$$

$$= \lim_{N\to\infty} \mathbb{P}\left\{ \sup_{n\geq N} \frac{S_n}{\sqrt{n}} \geq A \right\} \geq \lim_{N\to\infty} \mathbb{P}\left\{ \frac{S_N}{\sqrt{N}} \geq A \right\} > 0.$$

On the other hand, for all $m \geq 1$,

$$E = \left\{ \limsup_{n\to\infty} \frac{S_n}{\sqrt{n}} \geq A \right\} = \left\{ \limsup_{n\to\infty} \frac{X_m + \cdots + X_n}{\sqrt{n}} \geq A \right\} \in \mathcal{F}_m = \sigma(X_m, X_{m+1}, \ldots).$$

Thus, by Kolmogorov's Zero-One Law, $\mathbb{P}(E)$ is either 0, or 1. Since $\mathbb{P}(E) > 0$, we have

$$\mathbb{P}\left\{ \limsup_{n\to\infty} \frac{S_n}{\sqrt{n}} \geq A \right\} = 1$$

for all $A > 0$, implying the claim.

6. Let (X_n) be an i.i.d. sequence of random variables with

$$\mathbb{P}(X_n = 1) = 1/2 = \mathbb{P}(X_n = -1).$$

Let (Y_n) be a bounded sequence of random variables such that $\mathbb{P}(Y_n \neq X_n) \leq e^{-n}$. Show that

$$\frac{1}{n}\mathbb{E}(Y_1 + \cdots + Y_n)^2 \to 1 \text{ as } n \to \infty.$$

Solution: For $n \geq 0$, let A_n be the event

$$A_n = \{X_k \neq Y_k \text{ for some } k \geq n^{1/4}\}.$$

Then

$$\mathbb{P}(A_n) \le \sum_{k \ge n^{1/4}} \mathbb{P}(X_k \neq Y_k) \le \sum_{k \ge n^{1/4}} e^{-k} \le C_1 e^{-n^{1/4}}.$$

Now write $S_n = X_1 + \cdots + X_n$ and $T_n = Y_1 + \cdots + Y_n$. Using the Cauchy-Schwarz inequality,

$$\begin{aligned}
\mathbb{E}\left(\frac{S_n}{\sqrt{n}} - \frac{T_n}{\sqrt{n}}\right)^2 &= \frac{1}{n}\mathbb{E}(S_n - T_n)^2 \mathbf{1}_{A_n} + \frac{1}{n}\mathbb{E}(S_n - T_n)^2 \mathbf{1}_{A_n^c} \\
&\le \frac{1}{n}\sqrt{\mathbb{E}(S_n - T_n)^4 \mathbb{P}(A_n)} + \frac{1}{n}\mathbb{E}(S_n - T_n)^2 \mathbf{1}_{A_n^c} \\
&\le C_2 \frac{1}{n} n^2 e^{-(1/2)n^{1/4}} + \frac{1}{n}\mathbb{E}(S_n - T_n)^2 \mathbf{1}_{A_n^c}.
\end{aligned}$$

On the event A_n^c, one has $|S_n - T_n| \le C_3 n^{1/4}$, so the second term above is bounded by $\frac{1}{n} C_3^2 \sqrt{n}$. We obtain an overall bound of

$$C_2 n e^{-(1/2)n^{1/4}} + C_3^2/\sqrt{n},$$

which goes to 0 as $n \to \infty$. Therefore $S_n/\sqrt{n} - T_n/\sqrt{n} \to 0$ in L^2. Because $\|S_n/\sqrt{n}\|_2 = 1$ for all n, the triangle inequality gives

$$|\|T_n/\sqrt{n}\|_2 - 1| \le \frac{1}{\sqrt{n}}\|T_n - S_n\|_2 = \sqrt{\frac{1}{n}\mathbb{E}(S_n - T_n)^2} \to 0.$$

In other words, $\frac{1}{n}\mathbb{E}T_n^2 \to 1$.

7. Let F_n, F be distribution functions such that $F_n \to F$ weakly. If F is continuous, show that

$$\sup_{x \in \mathbb{R}} |F_n(x) - F(x)| \to 0.$$

Solution: Let $\epsilon > 0$. Because F is continuous, we may choose a finite collection of points $x_1, \ldots, x_K$ such that $F(x_i) = i\epsilon/3$. (Here K is chosen as $\lfloor 3/\epsilon \rfloor$.) Because $F_n \to F$ weakly, $F_n(x) \to F(x)$ at each continuity point x of F, and since F is continuous, $F_n(x) \to F(x)$ for all x. Thus we may choose N such that $n \ge N$ implies that $|F_n(x_i) - F(x_i)| < \epsilon/3$ for all $i = 1, \ldots, K$.

Now if $n \ge N$ and x is such that $x \in [x_i, x_{i+1}]$, one has

$$F(x_i) - \epsilon/3 < F_n(x_i) \le F_n(x) \le F_n(x_{i+1}) < F(x_{i+1}) + \epsilon/3,$$

and

$$F(x_i) \le F(x) \le F(x_{i+1}).$$

This means that both $F(x)$ and $F_n(x)$ are in the interval $(F(x_i)-\epsilon/3, F(x_{i+1})+\epsilon/3)$, and so

$$|F_n(x) - F(x)| < F(x_{i+1}) - F(x_i) + 2\epsilon/3 = \epsilon.$$

On the other hand, if $x < x_1$, $0 \le F_n(x) \le F_n(x_1) < F(x_1) + \epsilon/3$ and $0 \le F(x) \le F(x_1)$, giving $|F_n(x) - F(x)| < 2\epsilon/3$. Similarly, if $x > x_K$, then $F(x_K) - \epsilon/3 < F_n(x_K) \le F_n(x) \le 1$ and $F(x_K) \le F(x) \le 1$, giving $|F_n(x) - F(x)| < 2\epsilon/3$. Putting the three cases together, for $n \ge N$, $\sup_{x\in\mathbb{R}} |F_n(x) - F(x)| < \epsilon$.

8. Let (X_n) be an i.i.d. sequence of random variables. Show that $\mathbb{E}(X_1)^2 < \infty$ if and only if for every $c > 0$, $\mathbb{P}(|X_n| \ge c\sqrt{n}$ infinitely often$) = 0$.

Solution: Suppose first that $\mathbb{E}(X_1)^2 < \infty$. Then for any $c > 0$,

$$\sum_{n\ge0} \mathbb{P}(|X_n| \ge c\sqrt{n}) = \sum_{n\ge0} \mathbb{P}(X_n^2 \ge c^2 n) = \sum_{n\ge0} \mathbb{P}(X_1^2/c^2 \ge n).$$

We can write the right side using monotone convergence as

$$\mathbb{E} \sum_{n\ge0} \mathbf{1}_{\{X_1^2/c^2 \ge n\}} = \mathbb{E}\left(\left\lfloor \frac{X_1^2}{c^2} \right\rfloor + 1 \right) \le 1 + \frac{1}{c^2}\mathbb{E}X_1^2 < \infty.$$

So by the Borel-Cantelli lemma, $\mathbb{P}(|X_n| \ge c\sqrt{n}$ infinitely often$) = 0$.

Conversely, if $\mathbb{P}(|X_n| \ge c\sqrt{n}$ infinitely often$) = 0$, since the variables (X_n) are independent, these events are also independent, and so the Borel-Cantelli lemma (and reversing the above computation) gives

$$\infty > \sum_{n\ge0} \mathbb{P}(|X_n| \ge c\sqrt{n}) = \mathbb{E}\left(\left\lfloor \frac{X_1^2}{c^2} \right\rfloor + 1 \right) \ge \frac{1}{2}\mathbb{E}(X_1^2/c^2 + 1).$$

This implies that $\mathbb{E}X_1^2 < \infty$.

9. Find an example of a random variable X with a density function but whose characteristic function ϕ_X satisfies

$$\int_{-\infty}^{\infty} |\phi_X(t)| \, \mathrm{d}t = \infty.$$

Solution: Let X be exponential with mean 1. Then its characteristic function is

$$\phi_X(t) = \mathbb{E}e^{itX} = \int_0^\infty e^{itx}e^{-x}\ \mathrm{d}x = \frac{1}{1-it} = \frac{1+it}{1+t^2}$$

Therefore

$$|\phi_X(t)| = \frac{1}{1+t^2}\sqrt{1+t^2} = \frac{1}{\sqrt{1+t^2}} \geq \frac{1}{|t|},$$

and so

$$\int_{-\infty}^\infty |\phi_X(t)|\ \mathrm{d}t \geq \int_1^\infty \frac{\mathrm{d}t}{t} = \infty.$$

Probability Comprehensive Exam
Spring 2019

Student Number: ☐

Instructions: Complete 5 of the 9 problems, and **circle** their numbers below – the uncircled problems will **not** be graded.

1 2 3 4 5 6 7 8 9

Write **only on the front side** of the solution pages. A **complete solution** of a problem is preferable to partial progress on several problems.

1. Let X be a non-negative random variable, such that $0 < \mathbb{E}X < +\infty$, and let $0 < x < 1$. Show that
$$\mathbb{P}(X \geq x\mathbb{E}X) \geq (1-x)^2 \frac{(\mathbb{E}X)^2}{\mathbb{E}(X^2)}.$$

Solution:

If $\mathbb{E}(X^2) = +\infty$, then the inequality is trvially true. So without loss of generality, assume that $\mathbb{E}X^2 < +\infty$, in which case necessarily $\mathbb{E}X^2 > 0$, since $\mathbb{E}X > 0$. Next,
$$\mathbb{E}X = \mathbb{E}X\mathbf{1}_{X \geq x\mathbb{E}X} + \mathbb{E}X\mathbf{1}_{X < x\mathbb{E}X} \leq \mathbb{E}X\mathbf{1}_{X \geq x\mathbb{E}X} + x\mathbb{E}X.$$

Therefore,
$$(1-x)\mathbb{E}X \leq \mathbb{E}X\mathbf{1}_{X \geq x\mathbb{E}X} \leq \sqrt{\mathbb{P}(X \geq s\mathbb{E}X)}\sqrt{\mathbb{E}(X^2)},$$
by the Cauchy-Schwarz inequality. This proves the result.

2. If $(X_n)_{n\geq 1}$ is a sequence of random variables, then there exists a sequence $(c_n)_{n\geq 1}$ with $c_n \to \infty$, such that
$$\mathbb{P}(\lim_{n\to\infty} \frac{X_n}{c_n} = 0) = 1.$$

Solution: By Borel-Cantelli, it suffices to choose c_n such that
$$\sum_{n\geq 1} \mathbb{P}(|X_n| > \epsilon c_n) < \infty$$
for all choices of $\epsilon > 0$. We can choose first d_n such that $\mathbb{P}(|X_n| > d_n) < 1/2^n$ for each n. This is possible because for each n, $\mathbb{P}(|X_n| > \lambda) \xrightarrow[\lambda\to\infty]{} 0$, thus we can choose for each n a $\lambda_n > 0$ such that
$$\mathbb{P}(|X_n| > \lambda_n) < 1/2^n.$$

Using this, take $d_n = \lambda_n$. However to choose c_n we will take them such that $c_n = \max\{n, \max_{k=1,2,\ldots,n} d_k\}$. Clearly now, c_n is increasing to infinity and $c_n \geq d_n$. It remains to observe that for small $\epsilon > 0$
$$\sum_{n\geq 1} \mathbb{P}(|X_n| \geq \epsilon c_n) \leq \sum_{n\geq 1} \mathbb{P}(|X_n| \geq c_n) < \infty$$
and from this we get $|X_n|/c_n$ converges to 0 a.s.

3. Assume that $\{X_n\}_{n\geq 1}$ are random variables such that

    1. $E[X_n] = 0$ and $\mathbb{E}[X_n^2] \leq 1$ for any $n \geq 1$
    2. $\mathbb{E}[X_i X_j] \leq 0$ for any $i \neq j$.

    Show that for any sequence $\{a_n\}_{n\geq 1} \subset [1/2, 2]$,

$$\frac{a_1 X_1 + a_2 X_2 + \cdots + a_n X_n}{a_1 + a_2 + \cdots + a_n} \xrightarrow[n\to\infty]{\mathbb{P}} 0.$$

Solution: We use the standard proof of the weak law of large numbers for the case of finite variance. Denote $\bar{S}_n = \frac{a_1X_1+a_2X_2+\cdots+a_nX_n}{a_1+a_2+\cdots+a_n}$ and use Chebyshev's inequality to justify that

$$\mathbb{P}(|\bar{S}_n| \geq \epsilon) \leq \frac{\mathbb{E}[\bar{S}_n^2]}{\epsilon^2}.$$

Now,

$$\mathbb{E}[(\sum_{i=1}^n a_i X_i)^2] = \sum_{i,j=1}^n a_i^2 a_j^2 \mathbb{E}[X_i X_j] \leq \sum_{i=1}^n a_i^2 \mathbb{E}[X_i^2] \leq \sum_{i=1}^n a_i^2.$$

Thus we get that

$$\frac{\mathbb{E}[\bar{S}_n^2]}{\epsilon^2} \leq \frac{\sum_{i=1}^n a_i^2}{\epsilon^2 (\sum_{i=1}^n a_i)^2} \leq \frac{4n}{\epsilon^2 (n/2)^2} = \frac{16}{n\epsilon^2}.$$

Which proves the claim.

4. Let $(X_n)_{n\geq 1}$ be a sequence of non -negative uniformly integrable random variables such that, as $n \to +\infty$, $X_n \Rightarrow X$. Show that X is integrable and that $\lim_{n\to+\infty} \mathbb{E}X_n = \mathbb{E}X$.

Solution: By weak convergence, $\liminf_{n\to+\infty} \mathbb{P}(X_n > t) = \mathbb{P}(X > t)$, except possibly at countably many t, while by uniform integrability, the sequence $(\mathbb{E}X_n)_{n\geq 1}$ is bounded. Hence, by Fatou's Lemma,

$$\begin{aligned}\mathbb{E}X = \int_0^{+\infty} \mathbb{P}(X > t)dt &= \int_0^{+\infty} \liminf_n \mathbb{P}(X_n > t)dt \\ &\leq \liminf_n \int_0^{+\infty} \mathbb{P}(X_n > t)dt \\ &= \liminf_n \mathbb{E}X_n < +\infty.\end{aligned}$$

Next, for any $M > 0$,

$$\mathbb{E}X_n = \int_0^M \mathbb{P}(X_n > t)dt + \mathbb{E}X_n \mathbf{1}_{X_n \geq M} = \int_0^M \mathbb{P}(M > X_n > t)dt + \mathbb{E}X_n \mathbf{1}_{X_n \geq M},$$

and similarly,
$$\mathbb{E}X = \int_0^M \mathbb{P}(M > X > t)dt + \mathbb{E}X_n \mathbf{1}_{X \geq M}.$$
By uniform integrability, for each $\epsilon > 0$, there is an $M > 0$ such that the last terms in each one of the above equalities are less than ϵ. So, to conclude it is enough to show that as $n \to +\infty$, $\int_0^M \mathbb{P}(M > X_n > t)dt$ converges to $\int_0^M \mathbb{P}(M > X > t)dt$. But, since M can be chosen in such a way that $\mathbb{P}(X = M) = 0$, weak convergence and dominated convergence on $[0, M]$ give the conclusion.

5. If $X_1, X_2, \dots, X_n$ are iid exponential random variables with parameter 1, compute the almost sure limit of
$$\frac{1}{n}\sum_{i=1}^n e^{-X_k - 2X_{k+1} - 3X_{k+2}}$$
as n tends to infinity.

Solution: We split this according to
$$\begin{aligned}\frac{1}{n}\sum_{i=1}^n e^{-X_k - 2X_{k+1} - 3X_{k+2}} =& \frac{1}{n}\sum_{i=0}^{[n-3)/3]} e^{-X_{3i+1} - 2X_{3i+2} - 3X_{3i+3}} \\ &+ \frac{1}{n}\sum_{i=0}^{[(n-4)/3]} e^{-X_{3i+2} - 2X_{3i+3} - 3X_{3i+4}} \\ &+ \frac{1}{n}\sum_{i=0}^{[(n-5)/3]} e^{-X_{3i+3} - 2X_{3i+4} - 3X_{3i+5}} \\ &+ \frac{R_n}{n}\end{aligned}$$
where R_n is eventually a remainder which is certainly less than 2. Now using the strong law of large numbers, for each sum $\frac{1}{n}\sum_{i=0}^{[n-3)/3]} e^{-X_{3i+1} - 2X_{3i+2} - 3X_{3i+3}}$, $\frac{1}{n}\sum_{i=0}^{[(n-4)/3]} e^{-X_{3i+2} - 2X_{3i+3} - 3X_{3i+4}}$, $\frac{1}{n}\sum_{i=0}^{[(n-5)/3]} e^{-X_{3i+3} - 2X_{3i+4} - 3X_{3i+5}}$, we get that in almost sure sense, the limit is
$$\mathbb{E}[e^{X_1 - 2X_2 - 3X_3}] = \mathbb{E}[e^{-X_1}]\mathbb{E}[e^{-2X_2}]\mathbb{E}[e^{-3X_3}] = \int_0^1 e^{-2x}dx \int_0^1 e^{-3x}dx \int_0^1 e^{-4x}dx = \frac{1}{24}.$$

6. Assume $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space such that there exist $X_1, X_2 : \Omega \to \mathbb{R}$ two independent Bernoulli random variables such that $\mathbb{P}(X_i = 0) = \mathbb{P}(X_i = 1) = 1/2$. Show that Ω must have at least 4 elements.

   Give an example with Ω having 4 elements together with a sigma algebra such that on it we can define two independent Bernoulli as above.

   Can you generalize this?

Solution: Since $\mathcal{F}$ has at least 4 disjoint events, namely, $\{X_1 = 0, X_2 = 0\}$, $\{X_1 = 0, X_2 = 1\}$, $\{X_1 = 1, X_2 = 0\}$, $\{X_1 = 1, X_2 = 1\}$, it follows that in fact $\mathcal{F}$ contains more than 2^4 elements, for example, all sets which are disjoint unions of these four elements (also including the empty union) gives at least $2^4 = 16$ elements. Since $\mathcal{F} \subset \mathcal{P}(\Omega)$, it means that Ω must have at least 4 elements, otherwise $\mathcal{P}(\Omega)$ has at most 2^3 elements.

On $\Omega = \{0, 1\} \times \{0, 1\}$ and the sigma algebra of all subsets, we can define $X_1(\omega_1, \omega_2) = \omega_1$ and $X_2(\omega_1, \omega_2) = \omega_2$. This is the standard tensor product construction.

For a generalization, if we have n independent Bernoulli random variables then Ω must have at least 2^n elements. Indeed, we have 2^n disjoint subsets in $\mathcal{F}$ and thus $\mathcal{F}$ must have 2^{2^n} elements. This implies that Ω must have at least 2^n elements.

7. If X, Y are two random variables such that $X \geq Y$ and X, Y have the same distribution, then $X = Y$ almost surely.

Solution: We try to relate the cumulative functions. Thus

$$\mathbb{P}(X \leq x) = \mathbb{P}(Y \leq X \leq x) \leq \mathbb{P}(Y \leq x, X \leq x) = \mathbb{P}(Y \leq x) - \mathbb{P}(Y \leq x < X)$$

Thus, because $\mathbb{P}(Y \leq x) = \mathbb{P}(X \leq x)$, we get that $\mathbb{P}(Y \leq x < X) = 0$ for any choice of $x \in \mathbb{R}$. Finally,

$$\mathbb{P}(Y < X) \leq \sum_{r \text{ rational}} \mathbb{P}(Y \leq r < X) = 0.$$

Consequently, $\mathbb{P}(Y = X) = 1$.

8. Assume that $X_1, X_2, \ldots, X_n$ are iid with density $f(x) = \frac{2}{x^3}$ for $x \geq 1$ and 0 otherwise. Define
$$M_n = \frac{1}{n} \max\{X_1, \sqrt{2}X_2, \ldots, \sqrt{n}X_n\}.$$
   Show that X_n converges in distribution and find the limit.

Solution: We compute the cumulative function as for $x > 0$

$$F_{M_n}(x) = \mathbb{P}(X_1 \le nx, X_2 \le nx/\sqrt{2}, \ldots, X_n \le nx/\sqrt{n}) = \prod_{k=1}^{n} F_X(n^2 x/k).$$

Now the cumulative function of X is (for $x \ge 1$)

$$F_X(x) = \int_1^x \frac{2}{t^3} dt = 1 - 1/x^2.$$

Thus we have for $x > 0$ and large n, that

$$F_{M_n}(x) = \prod_{k=1}^{n} (1 - \frac{k}{n^2 x^2})$$

To compute the limit of this we take the log and use the fact that

$$\ln(1 - t) = -t + O(t^2)$$

for small, t, thus

$$\ln(F_{M_n}) = \sum_{k=1}^{n} \ln(1 - \frac{k}{n^2 x^2}) \approx -\frac{\sum_{k=1}^{n} k}{n^2 x^2} \xrightarrow[n\to\infty]{} -\frac{1}{2x^2}.$$

Therefore,

$$F_{M_n}(x) \xrightarrow[n\to\infty]{} F(x) = e^{-1/(2x^2)}, \text{ for } x > 0.$$

9. Let X be a finite mean random variable, let $\mathbf{F}$ be a σ-field and let G be a σ-field independent of $\sigma(\sigma(X), \mathbf{F})$. (As usual, $\sigma(X)$ is the σ-field generated by X and $\sigma(\sigma(X), \mathbf{F})$ is the σ-field generated by $\sigma(X)$ and $\mathbf{F}$.) Is it true or false that $\mathbb{E}(X|\sigma(\mathbf{F}, \mathbf{G})) = \mathbb{E}(X|\mathbf{F})$?

Solution: Yes it is true! Let $F \in \mathbf{F}$ and let $G \in \mathbf{G}$, then $F \cap G \in \sigma(\mathbf{F}, \mathbf{G})$ and using the very definition of conditional expectation as well as independence (twice) we get:

$$\begin{aligned}\mathbb{E}[\mathbb{E}(X|\sigma(\mathbf{F}, \mathbf{G}))\mathbf{1}_{F\cap G}] &= \mathbb{E}(X\mathbf{1}_{F\cap G}) = \mathbb{E}(X\mathbf{1}_F\mathbf{1}_G) = \mathbb{E}(X\mathbf{1}_F)\mathbb{E}\mathbf{1}_G \\ &= \mathbb{E}[\mathbb{E}(X|\mathbf{F})\mathbf{1}_F)]\mathbb{E}\mathbf{1}_G = \mathbb{E}[\mathbb{E}(X|\mathbf{F})\mathbf{1}_F)\mathbf{1}_G] = \mathbb{E}[\mathbb{E}(X|\mathbf{F})\mathbf{1}_{F\cap G}].\end{aligned}$$

Therefore $\mathbb{E}(X|\sigma(\mathbf{F}, \mathbf{G}))$ and $\mathbb{E}(X|\mathbf{F})$ agree on a π-system generating $\sigma(\mathbf{F}, \mathbf{G})$. Now, let μ_1 and μ_2 be respectively defined via $\mu_1(A) = \mathbb{E}[\mathbb{E}(X|\sigma(\mathbf{F}, \mathbf{G}))\mathbf{1}_A]$ and $\mu_2(A) =$

$\mathbb{E}[\mathbb{E}(X|\mathbf{F})\mathbf{1}_A]$. Since $\mathbb{E}|X| < +\infty$, then μ_1 and μ_2 are finite measures which agree on a π-system generating $\sigma(\mathbf{F}, \mathbf{G})$, so they must agree on $\sigma(\mathbf{F}, \mathbf{G})$. Finally, by uniqueness,

$$\mathbb{E}(X|\sigma(\mathbf{F}, \mathbf{G})) = \mathbb{E}(X|\mathbf{F}).$$

This proves the result. (Above, instead of μ_1 and μ_2, one could also consider positive measures by looking at the positive and the negative part of X).